Crop cycles and hierarchy: the agro-ecological origins of the state*

Preliminary draft

Matteo Sestito†

October 12, 2023

Abstract

Subject to exoduses, internal rebellions and diseases, states have historically developed only under very particular agro-ecological circumstances. This paper advances and empirically validates a new perspective on state formation, helping to understand their paucity and uneven development across the globe throughout the pre-industrial era. I posit that spatially asynchronous agricultural calendars mandated for a scattered productive system that could not be easily taxed and harnessed by centralised governments. Using data from the Ethnographic Atlas, I provide evidence that the heterogeneity of agricultural growing seasons was a crucial barrier to state centralisation. This holds true when controlling for a wide range of alternative determinants of state-building. The use of potential, rather than observed, agro-ecological data, as well as various robustness tests, give credit to an interpretation of the results beyond the mere correlation. Additional evidence on 19^{th} century taxation in India and China, sheds light on the precise mechanisms whereby crop cycle heterogeneity hindered political centralisation.

JEL codes: D02, H10, N50, O43

*I would like to thank my advisor, Nicolas Berman, for his patience and precious attention. Comments by Anatole Cheysson, Yannick Dupraz, Luigi Pascali, and David Schönholzer have been of great help. This paper has benefited from useful comments by seminar participants at Aix-Marseille School of Economics, Pompeu Fabra University, Barcelona School of Economics PhD Jamboree 2022, Rare Voices in Economics conference 2022, ASREC European conference 2022, ASSET conference 2022, YSI-EHES 2023 webinar. It goes without saying that any error in the work is my own responsibility.

†PhD candidate at Aix-Marseille School of Economics

1 Introduction

> *I am Xerxes, the great king, the king of kings, the king of the provinces with many tongues, the king of this great earth far and near, son of king Darius the Achaemenian*
>
> Xerxes I

Despite the grandiose tone of the Achaemenid propaganda,[1] Xerxes' claims of a global monarchy were rather far-fetched when seen from a global perspective. During his reign (486-465 BCE), the Persian empire extended far and wide, from the Turkish coasts in the West to the Indus Valley in the East. Other kingdoms dotted the eastern Mediterranean basin as well as south-western and eastern Asia; centralised governments were consolidating in Meso-America and the northern portions of the Peruvian coast too. Yet, despite their sumptuous palaces and temples, states remained relatively isolated polities, confined to particular agro-ecological niches and submerged into a world made of tribal confederations, petty chiefdoms, acephalous forager bands, dispersed horticulturalist villages and nomadic maritime communities. While nowadays virtually all of humankind organise itself into states − that is, under centralised forms of government involving a bureaucratic apparatus with several layers of authority −[2] for much of human history, people have experimented and played with a great variety of social arrangements. It is, arguably, not until the 17^{th} century that the majority of humankind came to live under the yoke of a centralised government (Scott 2017).

Even on the eve of the industrial revolution, large swaths of the planet had never seen a state bureaucrat or had, at most, experienced weakly centralised forms of government. These lacks would often prove to be fatal. It was the complex of "guns, germs and steel" manned by Eurasian states that would eventually conquer, colonise and exterminate the various non-state polities spread across the rest of the globe (Diamond 1997). Understanding, why states emerged only in some areas of the world is thus a question with profound historic ramifications. It

1. The sentence comes from an trilingual inscription in Elamite, Old Persian and Babylonian at a fortress near the Lake of Van in eastern Anatolia (Dusinberre 2013, pp. 50-54).

2. There is not universally accepted definition of what exactly is a state. Jurists and political scientists tend to follow the Weberian tradition, describing states as governments that, enjoying a monopoly over the use of force, control a permanent population within a fixed territory (Mullerson 1993; Ikenberry 2011; Shaw 2014, ch. 5). As these features − monopoly of violence and territorial sovereignty − are notably lacking in many polities of the pre-industrial period, anthropologists and archaeologists have opted for more general definitions, identifying states on the base of the presence of a multi-tier political organisation, where decision-making is articulated over different administrative levels (Claessen 2004, Diamond 1997, Grinin 2004, Redmond and Spencer 2012, Trigger 2003, ch. 10).

has also direct economic consequences, as many studies have shown (Bockstette et al. 2002). For example, pre-colonial state centralisation in Sub-Saharan Africa has been associated to lower children mortality (Gennaioli and Rainer 2007) and higher economic development (Michalopoulos and Papaioannou 2013).

This work addresses the question on the origins of the state from a new angle. It tries to understand the spatial distribution of early states by putting the latter in relation to the concentration of economic activity. Where this is spatially and temporally concentrated, centralised governments can arise and endure because they find easier to organise and tax production. In the Malthusian pre-industrial world, economic activity primarily consisted of agricultural production, and taxation was indeed mainly conducted through expropriation of part of the harvest and the imposition of corvée labour.[3] Thus, in areas where neighbouring fields followed different agricultural calendars: (*i*) taxation was inherently difficult because it required the deployment of tax collectors at different moments of the year and because production could be more easily concealed; (*ii*) the extraction of labour and military services from the local population was fragmentary and ineffective inasmuch as farmers could not be mobilised all at once.

While the above discussion might seem rather abstract, the control of the agricultural cycle has historically been of great concerns to state rulers. As noted by Scott (2017, p. 133): "Archaic states endeavoured, whenever possible, to mandate a planting time for a given district." For example, in archaic China fields were forcibly irrigated at the same time so as to impose a common growing season to all rice cultivators.

These policies shall be framed within the greater homogenisation effort sponsored by many agrarian states of the pre-industrial era. Indeed, the rise of centralised governments is almost everywhere accompanied by parallel attempts to standardise a wide host of vernacular practices so as to measure, predict and eventually appropriate the resources at their disposal. Hence, the imposition of standard weights, metrics, and units of account had the twofold effect of facilitating economic transactions and of making production more transparent to state bureaucrats. For example, during the late 3^{rd} millennium BCE, the introduction in Mesopotamia of the "labourer-day" unit of account made economic performances both comparable across different sectors and measurable in their own right, thus

3. For example, archaic China levied a 11% harvest tax and requested one month of corvée labour. Pharaoh Egypt had similar tax rates on the agricultural production, albeit demanding longer labour services (Schönholzer 2020). The highly centralised Neo-Sumerian Empire of the Third Dynasty of Ur demanded tributes to its subject cities as high as 48% of the barley harvest, beyond assigning several corvée tasks in the fields (Adams 2007 & 2008). Even in more recent periods, in-kind payment remained central to tax collection. Half of Tokugawa Japan's tax receipts were, for example, in rice until the late 18^{th} century (Sato 1990, p. 44).

allowing the imposition of precise working duties (Carmona and Ezzamel 2007). But homogenisation went well beyond the mere economic sphere, touching even upon the sacred, with religious practices and architecture becoming increasingly standardised as states asserted themselves.[4] The rise of Monte Alban in the Mexican highland (500-300 BCE) or of the Mayan city-states in the Yucatan peninsula (250-500 CE), were, for example, characterised by the parallel emergence of standardised two-room temples (Redmond and Spencer 2012).

The homogenisation of the crop calendar should be thus read through these lenses and understood as a further attempt to homogenise the forms and tempo of social life. Indeed, much of the economic and religious life of agrarian communities rotates around the growing season of their main staple crop. This mandates for concrete working schedules – cleaning fields, sowing, planting, tending, harvesting – as well as for particular harvest celebrations and fertility gods. In short, heterogeneous growing seasons translated into heterogeneous social preferences, which ultimately represented a barrier to the homogenisation effort of early states.

The present paper brings this hypothesis to the data, exploring the impact of growing period heterogeneity onto state centralisation in the pre-industrial world. Measures of political hierarchy from the Ethnographic Atlas are combined with agro-ecological characteristics as retrieved from the Food and Agricultural Organisation.[5] Various other spatial databases are also used to gather geographical, climatic and socio-economic information. This allows for the construction of a sample spanning the whole globe, and including more than 1200 pre-industrial societies.

The empirical exercise is fraught with difficulties. As the above discussion should have made us well aware, there are obvious risks of reverse causality: crop cycle homogeneity facilitated the emergence of the state, which in turn endeavoured to establish agro-ecological settings marked by synchronous agricultural calendars. To circumvent these risks, I employ data on the potential, rather than actual, productivity of crops. Thanks to this largely exogenous measure, this work establishes a strong negative correlation between crop cycle heterogeneity and state centralisation. A wide set of controls ensures that the correlation is not driven by some omitted variable, ruling out that the effect of growing period heterogeneity is mediated by other factors traditionally associated to state-building, such as: the productive advantage of storable crops over perishable ones (Mayshar et al. 2022, Scott 2017); the degree to which societies are circumscribed by inhospitable lands

4. See Flannery (1998), Diamond (1997, p. 280).

5. The Ethnographic Atlas is an anthropological database widely used in social sciences. See Michalopoulos and Papaioannou (2013), Fenske (2014), Enke (2019), and Mayshar et al. (2022) for works in economics employing, *inter alia*, the same variable on state centralisation here used. See Kelly (2007) for a review of some articles in anthropology using this dataset.

(Carneiro 1970, Mayoral and Olsson 2019, Schönholzer 2020); the easiness at conducting trade (Algaze 2009, Fenske 2014, Litina 2014, Tedeschi 2021); the presence of waterways and irrigation canals (Allen et al. 2020). Further evidence on the detrimental effect of crop cycle heterogeneity on political centralisation is provided with the aid of two additional independent datasets on 19^{th} century taxation in Qing China and colonial India. For both polities, land tax revenues are significantly lower in districts with more diverse agricultural calendars. Overall, the analysis enriches our understanding on the origins of the state, shedding further lights on its agro-ecological limits.

The rest of the paper is organised as follows. The next section provides a conceptual framework to understand state-building in the Malthusian era and it illustrates the channels whereby crop cycle heterogeneity hindered political centralisation. Section 3 discusses the data sources used in the empirical analysis, which is presented at section 4. A final section concludes.

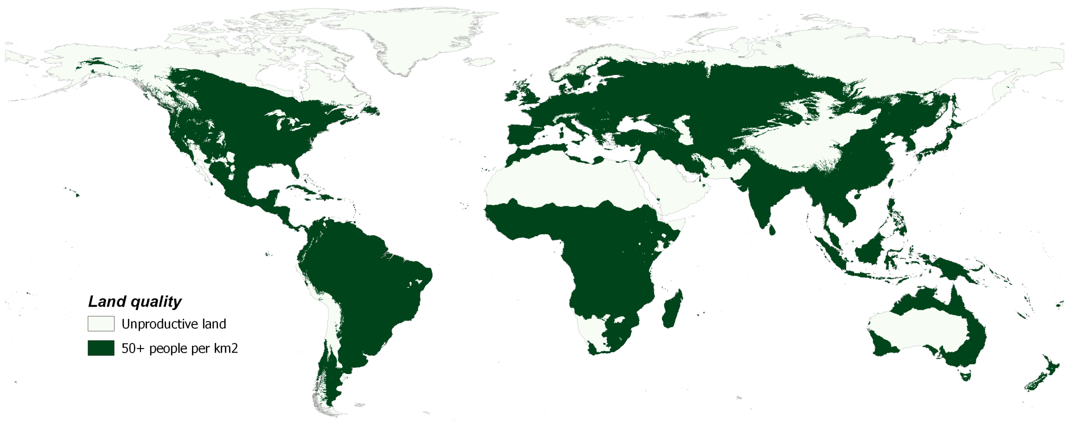
2 Conceptual framework

The origins of the state have stimulated the intellectual curiosity of scholars across the whole spectrum of social sciences. Particular attention has been spent on those supposed cases of pristine state emergence, that is, those instances where a centralised government emerged without any interference from other peer polities. The inherent problems in the notion of "pristine", not to say on that of "state", are now largely recognised and the attention has shifted towards more general treatments of state-building.[6] The traditional and arguably more popular view considers states as the more or less natural outcome of the domestication processes initiated during the Neolithic.[7] Where societies adopted intensive farming techniques and land was fertile enough, population densities rose spectacularly, eventually paving the way to the emergence of state-like bureaucratic structures. Large societies, whose population numbers in the hundreds, are mainly composed of people who are strangers to themselves, thus requiring impersonal law and institutions to manage conflict, redistribution, or reach other types of communal decisions (Diamond 1997). As testified by the emergence of most of the early

6. On this point see: Claessen (2004), Graeber and Wengrow (2021), Grinin (2004), Possehl (1998).

7. To put it as Gat (2006, p. 232): "State evolution was the almost 'necessary' culmination and fruition of processes set in motion by the transition to and growth of agriculture - at least where the right conditions were present." Similar arguments can be found also in Diamond (1997) and to a lesser extent in Harari (2014). Graeber and Wengrow (2021) discuss the poignancy of the idea even among specialists.

Figure 1: Land productivity



Notes: The underlying land productivity data are taken from the FAO-GAEZ dataset, under the low-input & rain-fed specification. The daily caloric intake used is 2000 calories per person.

states on rich alluvial soils near centres of domestication, there is, indeed, a broad grain of truth in this.

Yet, the argument is somehow too general and has not much explanatory power. Figure 1 shows the portions of the globe that can support at least 50 people per square kilometre. The threshold has been purportedly set at a quite elevated level, higher than actual densities historically achieved in the pre-industrial world.[8] As clear from the figure, apart from some deserts and mountainous ranges, much of our planet is productive enough to sustain dense populations: land productivity by itself can not explain the rise of states.

The traditional view has received a lot of criticism also for its linear evolutionary flavour, whereby mankind is seen as progressing from simple nomadic hunter-gathering societies to settled farming communities and eventually to kingdoms and civilisations. In the journey, so the argument goes, we bartered equality (Marxist primitive communism) for peace and order. Yet, for how appealing in its simplicity, this account does not match well with historical reality. For example, we now have ample evidence that sedentism predated farming in many settings, such as in the Levant and southeastern Turkey, Jomon Japan, Sudan, and perhaps central Mexico and the northern Andes too (Bellwood 2004). The construction of cities and the rise of political hierarchies, usually hailed as hallmarks of states and civilisations, are also found under previously unexpected circumstances. The

8. To give some context, Renaissance Italy had a population density of about 30 people per km^2, the highest in Europe at the time. India and Mexico in 1500 had population densities of 44 and 13 people per km^2. Population estimates are collected on the website Our World in Data, freely accessible at: https://ourworldindata.org/grapher/population-density?time=1500.

monumental architecture at Goebli Tepe and Poverty Point was, for example, the making of foraging, rather than farming, societies (Graeber and Wengrow 2021). Similarly, the hunter-gathering aristocracies of the north-western coast of North America, with their slaves and retinues, further disprove the necessity of farming in fostering social inequality (Kelly 2007, ch. 9).

But perhaps the most severe flaw of the land productivity theory is its inability to explain the widespread opposition to states and central governments. Instead of being an irresistible force which everyone hailed to join, the imposition of vertical structures of power was usually hotly resisted. For any successful state-building attempt, there are many more that failed and were often lost in the dust of history: for any Alaric, there are several Maroboduus and Arminius whose royal aspirations were blocked by competing elites and commoners alike (Gat 2006). Even once states managed to assert themselves, they were extremely fragile entities constantly menaced by popular revolt and flee. The historical records is replenished of examples of people escaping what was perceived as a too onerous tax burden. From the Semang of Malaysia escaping the oppressive rule of Malay and colonial authorities (Scott 2009), to the Guayaki of Paraguay escaping the colonial reducciones and slave raids (Clastres 1987), many people preferred a life in the wood, distant from the "civilised" palaces of early agrarian states. This is not to deny that people moved also in the other direction, abandoning the barbarian frontier when allured by the economic and religious power of early states. Yet, from the perspective of central authorities, the constant fear of a people haemorrhage was very much present and shall be taken into account when discussing the origins of the state. As put it bluntly by Scott (2017, p. 30): "The great walls of China were built as much to keep Chinese taxpayers in as to keep the barbarians out."

As the link between land productivity and political hierarchy comes increasingly under scrutiny, state-building studies have tended to emphasise those conditions facilitating the control and taxation of local population. Particular attention has been spent on those factors limiting the possibilities of outmigration. The idea here being that states emerged where people could not easily flee away from their yoke (Carneiro 1970, Dickson 1987). Empirical evidence on this point comes from Schönholzer (2020), who find that pristine state formation is associated to land circumscription, measured as the differential in land productivity between a zone and the neighbouring areas. Similar evidence is available for ancient Egypt, where state power is correlated to positive productive shocks in the core Egyptian territories and negative agricultural shocks in its periphery (Mayoral and Olsson 2019).

Alongside migration possibilities, another element positively influencing the tax base of (would-be) states is the presence of patchy, regular, and appropriable resources (Smith, Mulder, et al. 2010). The presence of storable and predictable agricultural surpluses is particularly relevant in discussions on the consolidation of the first states and is often considered almost a necessary condition for their emergence (Scott 2017). Empirical evidence on this point comes from Mayshar et al. (2022), who find that archaic and pre-modern states were more centralised where the production of storable crops such as cereals enjoyed advantages over the cultivation of more perishable roots and tubers.

Crop cycles represent a new, relatively neglected, agro-ecological constraint on the fiscal capacity of pre-industrial polities.[9] In particular, the fragmentation of the agricultural calendar might have hindered political centralisation through two main channels. First, expropriation of the harvest is harder in settings with heterogeneous crop cycles inasmuch as: tax collectors need be to dispatched several times per year to a same area; production is opaque and not easily assessable. Second, exploitation of the local labour force, either for military or working services, is rather inefficient because farmers can not be mobilised all at once.

To fix ideas, consider two regions: a first region A, where adjoining rural areas have crops with staggered crop cycles; and a second region B, whose cultivars are planted and harvested at the same moment of the year. In order to collect taxes from A, government officials would have to travel to the region several times per year, at great logistical and economic cost. The taxation of B is, instead, inherently easier, being a one-stop affair: when government officials are dispatched there, they can appropriate the bulk of the total yearly production. Moreover, synchronous crop cycles make production more transparent: as cultivars are planted and mature at the same time, they are easily observed and registered in state bureaucracy, eventually helping central authorities in censusing agricultural production. Furthermore, as the bulk of farmers in B is freed from its agricultural duties at the same period, a huge reservoir of labour can be harnessed by government in order to wage war and build public works.

It is hard not to overstate the importance of these constraints on state capacity. Making production legible and assessable has historically been a major concern of central rulers.[10] The measurement effort conducted by agrarian states becomes tangible in the countless inventory lists, accounting books, cadastral maps and population censuses found in the archaeological and historical record (Carmona and Ezzamel 2007). It is perhaps no chance that writing is found among all state

9. Importantly, as it will be shown in Section 4, the impact of heterogeneous growing seasons on political centralisation is largely orthogonal to land circumscription and the so-called cereal advantage.

10. On the importance of production transparency for state-building outcomes see: Mayshar et al. (2017), Sánchez De La Sierra (2020).

societies and that, before chanting the feasts of gods and heroes, it remained for long confined to the economic and accounting sphere (Rahmstorf 2012, Harari 2014).

If accessibility and transparency of production were of great importance for pre-industrial polities, the seasonal exploitation of a huge reservoir of labour was of not lesser consequence. As the *common* agricultural period came to end, farmers were employed in the construction of pompous monuments aimed at glorifying rulers and legitimising state power. The impressive Egyptian pyramids were, for example, built during the flooding months of the Nile, when farming activity was on halt. Similarly, from Hawaii to ancient Mesopotamia, the maintenance of irrigation canals and dams was a communal effort, requiring the coordination of labour forces (Diamond 1997, Allen et al. 2020). In some cases, the very establishment of state structures seem to be related to massive labour investments, as during the High Middle Ages in central Poland (Lozny 2004). Finally, also warfare followed the rhythm of agriculture, with conscription and military campaigns beginning after the main harvest had been collected (Scott 2009, Trigger 2003).

Heterogeneous crop cycles thus hampered state capacity, understood as the ability of central governments to raise tax and command labour. Moreover, the structure of the agricultural calendar had important repercussions over local social preferences. In the Malthusian era, the rhythm of communal life was mandated by the agricultural growing season. Working schedules revolved around the needs of the main staple crop: there was a time for sowing and planting, one for tending the plants, and eventually the period of the harvest. Most of the phases in the crop cycle were accompanied by specific religious rituals and other cultural peculiarities: fertility gods, harvest rites, ceremonial ploughing, and seasonal festivals. In places marked by a homogenous agricultural calendar, there was much less room for diverging cultural idiosyncrasies: "the uniformity in the field [...] produced a social and cultural uniformity expressed in family structure, the value of child labour and fertility, diet, building styles, agricultural ritual, and market exchange."[11] Hence, heterogeneous growing seasons might have also resulted into a mosaic of diverse habits, or more broadly, into heterogeneous social norms. It is not difficult to imagine how this very disparate patchwork of cultural practices might have represented a barrier to the formation of centralised governments, with each community guarding jealously its traditions against the common religious and economic standards associated to the rise of states.

Beyond crop cycles, there are of course many other factors that have been

11. The quote is from Scott (2009, p. 75), who actually refers to uniformity in both the timing and type of agricultural production. The same logic, however, applies to homogeneity in the tempo of production.

proposed as causes and triggers in the process of (early) state-building, ranging from warfare (Gat 2006, Spencer 2010) to trade (Algaze 2009, Fenske 2014, Litina 2014, Tedeschi 2021) and irrigation (Allen et al. 2020).[12] Any attempt to rank by importance the various forces leading to political centralisation is bound to be unsatisfactory. Each state society has its own unique story, involving a different sets of triggers and causes leading to the adoption of a centralised bureaucratic government. Yet, some general patterns can be discerned and some negative conclusions can be advanced. As noted by a leading scholar of early states: "It is surely striking that virtually all classical states were based on grain, including millets. History records no cassava states, no sago, yam, taro, plantain, breadfruit, or sweet potato states."[13] This is not to say that cereals caused political hierarchy. Maize cultivation was, after all, well known by the indigenous people of the northeastern and mid-western woodlands of North America (Graeber and Wengrow 2021). Cereals simply represented an efficient medium of taxation: where their cultivation was preferred to other perishable crops *and*, as we shall see, their growing season was homogenous, a central authority could sustain itself by extolling a tribute from the local population. Absent these ecological conditions, states could potentially emerge anyway; for example, in virtue of an incredibly strong and fervent ideology. Yet, episodes of this type are bound to be short-lived. The countless prophetic movements of the Lahu and Karen of mainland south-east Asia provide clear examples: these experiments of supra village governance and alliances faded away as soon as their charismatic momentum died out (Scott 2009, ch. 8). Hence the importance of some structural agro-ecological factors necessary to durably sustain taxation and extraction on a large scale. As the rest of the paper will substantiate, historically, the homogeneity of crop cycles has figured prominently among these.

3 Data

3.1 Dependent variable

Data on state centralisation is taken from the Ethnographic Atlas (EA), a dataset largely used in both the economic and anthropological literature.[14] The EA contains information on 1249 pre-industrial societies observed after 1500 CE.[15] The

12. For space reasons, the discussion of these factors is relegated to Appendix A.

13. Scott (2017, p. 21). A remarkable exception to this broad empirical regularity is the Yoruba society, with its city-state system relying mostly on yam cultivation before the spread of New World crops (Trigger 2003, p. 285).

14. See footnote 5 for a list of recent works employing the Atlas.

15. The original database includes 1265 societies. Eight observations have then been dropped because relative to pre-Columbian times; eight societies have been excluded from the analysis because the year of observation is missing.

sample has a good coverage of North America and Africa, while reporting few European societies.[16]

The EA variable used to capture state centralisation measures the levels of jurisdictional hierarchy above the local community. This ordered variable is the standard measure of political complexity used in the literature. It ranges from 0 to 4 and has been coded without considering organizations not held to be legitimate, such as imposed colonial regimes (Murdock 1967, p. 52). A value of 0 indicates acephalous societies organised in autonomous villages. The presence of 1 jurisdictional level describes societies where local communities are directly politically subordinated to some elite, as in petty chiefdoms and Melanesian tribes ruled by "big-men". Higher scores correspond to large chiefdoms and states, that is, societies endowed with a multi-layered administrative apparatus at their head. Examples of societies without any centralised political organisation are: the Comanche of the Southern Plains in US, the Herero pastoralists of Southern Africa, the Semang of the Malay peninsula, and the Amazonian Yanomamo. At the other extreme there are polities with four levels of jurisdictional hierarchy such as: the Siamese state in modern-day Thailand, the Punjabi people inhabiting the homonymous region between Pakistan and India, the Bubi of Equatorial Guinea, and the Kafa of Ethiopia.

The majority of the sample is, however, represented by acephalous societies. Figure 2 reports the histogram of the state centralisation variable: more than 70% of the societies have at most one level of political hierarchy. Figure 3 gives a visual representation of the societies in the Atlas, employing the ethnic maps assembled by Fenske (2014).[17] The ethnic polygons are shaded on the base of each society's centralisation level: stateless societies are particularly common in the Americas, while Eurasia shows deeper political hierarchies.

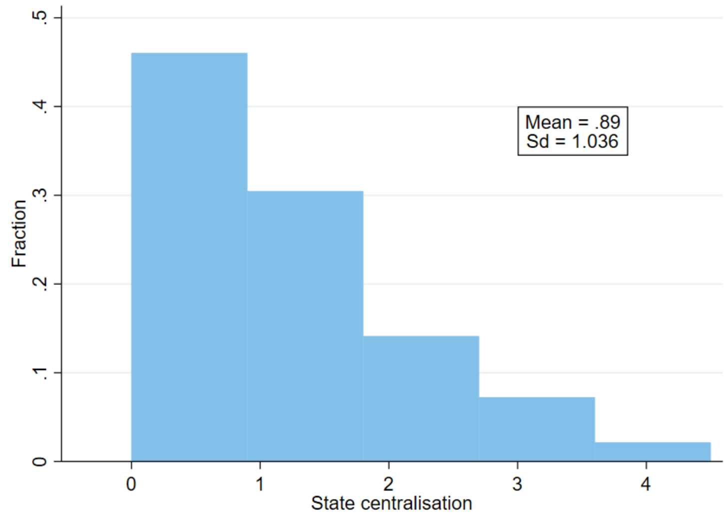
It shall be stressed that the societies of the Ethnographic Atlas have been sampled mostly towards the late pre-industrial era, with the focal year of their observation referring predominantly to the late 19^{th} century. A plausible concern is thus the idiosyncratic nature of these polities. For example, was the Kafa Kingdom an historical accident observed only in 1890 when it was sampled? Or its presence is a symptom of a longer state tradition? As any expert of Ethiopian history would know, the kingdom dates back to the late 15^{th} century and, more importantly, emerged in a region where states vied for power at least since the emergence of the Axum Empire in the 1^{st} century BCE (Butzer 2012).

Beyond the political vagaries of the Horn of Africa, state institutions are relatively persistent throughout history. Figure 4 shows the autocorrelation through

16. All the empirical analysis exploits within-continent variation and results are robust to the sequential exclusion of each continent.

17. The polygons have been developed by Fenske (2014) upon consultation of various sources, ranging from historic maps to current administrative boundaries.

Figure 2: State centralisation - Histogram



time of an index of state centralisation developed by Borcan et al. (2018). The authors, extending the work of Bockstette et al. (2002), assign a measure of state presence to each present-day country at intervals of 50 years, from archaic to present times. The autocorrelation of past statehood with respect to statehood in 1800 CE is positive and increasing over time, peaking at almost 0.6 for statehood in 1500 CE. Importantly, state persistence is observed only in the pre-industrial period: state presence today is not predicted by past statehood, with autocorrelation coefficients hoovering around zero. While the projection of contemporary boundaries back in time is inherently problematic, the exercise gives credit to the idea that some relatively permanent constraints determined the emergence of states only in some specific areas of the world.

Figure 3: Pre-industrial state centralisation

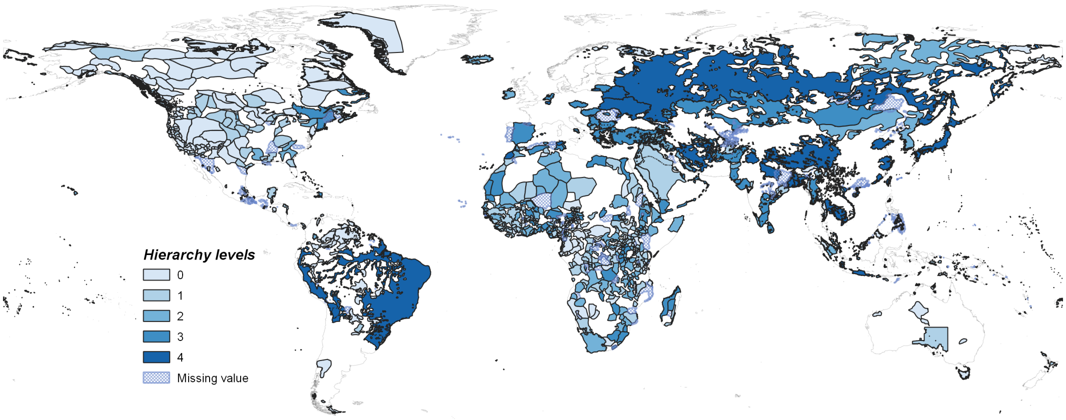
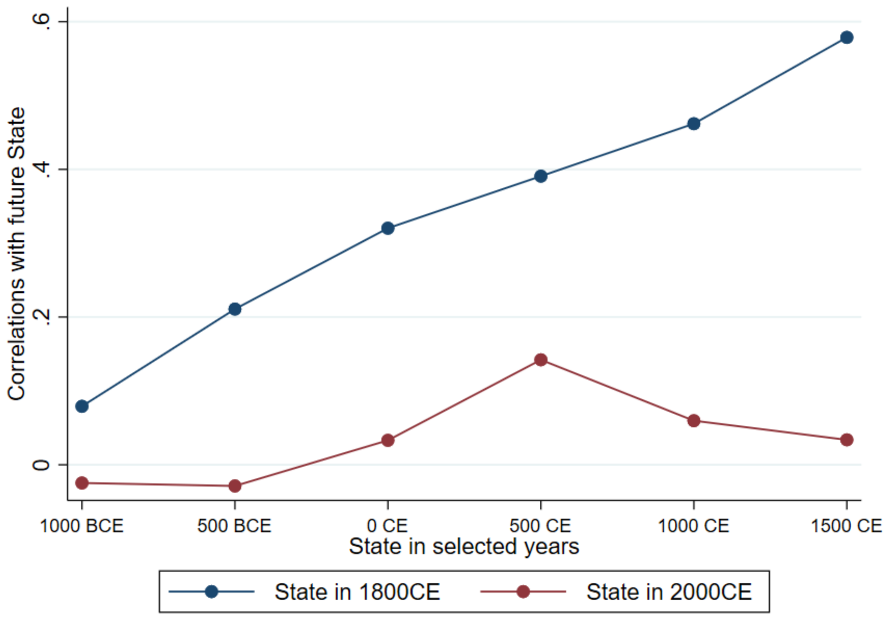


Figure 4: Temporal autocorrelation of State index



3.2 Growing period heterogeneity

The measure of crop cycle heterogeneity is built using the Global Agro-ecological Zones (GAEZ) dataset of the Food and Agriculture Organisation (FAO). GAEZ reports crop yields and growing seasons for a set of 40 edible crops. The data span the whole globe and are in raster format, with pixels at the 0.083° resolution ($\sim 81km^2$ at the equator). The data refer to *potential*, rather than observed, productivity and growing seasons, and they are computed under consideration of agro-climatic constraints. Importantly, these constraints exclude agro-edaphic factors, such as soil salinisation, that are directly affected by human intensive farming techniques. GAEZ measures are thus largely exogenous to human activity, lessening risks of reverse causality. Furthermore, among the various specifications, I employ FAO estimates based on farming practices relying on low inputs and rain-fed water supply. These conditions are arguably independent of human intervention and better describe pre-industrial agricultural settings.[18]

To capture the extent to which a given area has an heterogeneous crop cycle, I compute the fraction of the year whereby the main cultivar in pixel p does not grow at the same time as the major cultivars in the surrounding pixels. Define GP_p as the day-unit set describing the growing period of the most productive crop in pixel p.[19] Then, two cells p and k have different crop cycles when their main growing seasons do not overlap too much throughout the year. Define thus:

$$GP_{p,k}^{het} = 1 - \frac{|GP_p \cap GP_k|}{365}$$

The measure has a simple interpretation as the fraction of the year where the main crops of p and k do not grow together. For example, $GP_{p,k}^{het} = 1$ indicates the maximum possible heterogeneity, attained when the main growing period of the two areas are completely disjoint (*i.e.* $GP_p \cap GP_k = \emptyset$). Similarly, $GP_{p,k}^{het} = 0.5$ indicates that for half of the year the two cells share the same agricultural calendar.

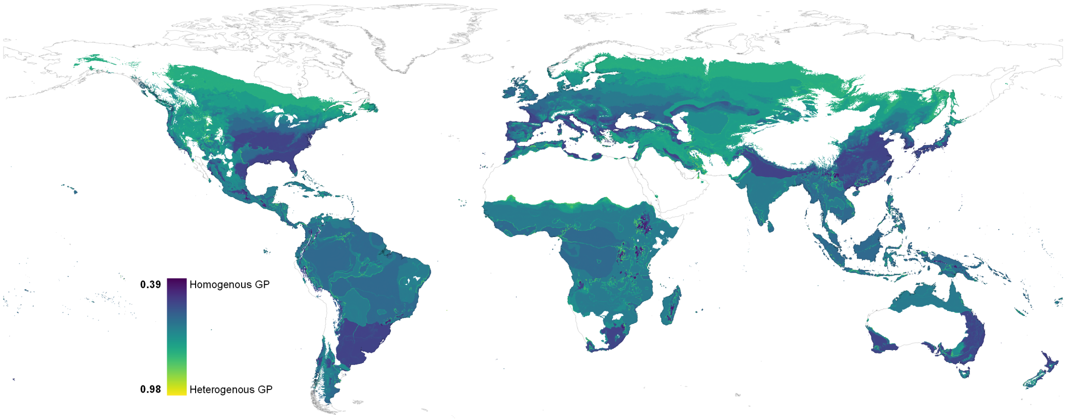
To get a variable at the pixel level, this bilateral heterogeneity measure is aggregated over the cells surrounding each given pixel. Define N_p as the set of pixels within the neighborhood of p, then growing period heterogeneity at the pixel level is defined as:

$$GP_p^{het} = \frac{1}{|N_p|} \sum_{k \in N_p} GP_{p,k}^{het}$$

18. This specification of GAEZ data − exclusion of agro-edaphic contraints, low-input & rain-fed farming regime − is the one most commonly used in the literature interpolating FAO data to the pre-industrial period. See for example: Galor and Özak (2016), Mayshar et al. (2022).

19. Productivity is measured in calories, with GAEZ ton/hectare data transformed into calorie/hectare using FAO nutritional tables. Details on the caloric content of each crop are given in Table A1.

Figure 5: Growing period heterogeneity



The baseline analysis employs the 8-pixel neighbourhood (*i.e.* $|N_p| = 8$), but in robustness exercises alternative neighbourhood sizes are checked.

Figure 5 shows the global distribution of GP_p^{het}. Already at first glance, it can be seen that some areas traditionally related to state presence (*e.g.* East Asia) display more homogenous crop cycles relatively to the rest of their continent. The measure is then aggregated at the society level by averaging it across all pixels belonging to the ethnic polygon of a given society.[20] Figure 6 reports the distribution of the growing period heterogeneity variable at the societal level, showing that much of the variability is concentrated around intermediate levels of heterogeneity.

It shall be stressed that GAEZ data refer to the second half of the 20^{th} century. However, as noted by Nunn and Qian (2011, p. 611), who use the dataset to capture agricultural suitability in the 18^{th} century, GAEZ "measures should be good proxies for historical conditions because they are primarily based on climatic characteristics such as temperature, humidity, length of days, sunlight, and rainfall that have not changed significantly [over the last few centuries]". The authors also provide evidence that GAEZ suitability for potato well correlates with actual potato production in 1900.

As a further test of the validity of GAEZ data to describe historical conditions prevalent in the late 19^{th} century, Table 1 displays estimates from a regression of the subsistence practices reported in the EA on an index of land quality, as proxied by the caloric yield of the most productive crop in each pixel. Dependence on agriculture and farming intensity are strongly positively correlated to land quality.

20. In robustness exercises, I check alternative geographical representations of the Ethnographic Atlas societies.

Figure 6: Growing period heterogeneity - Histogram of EA societies

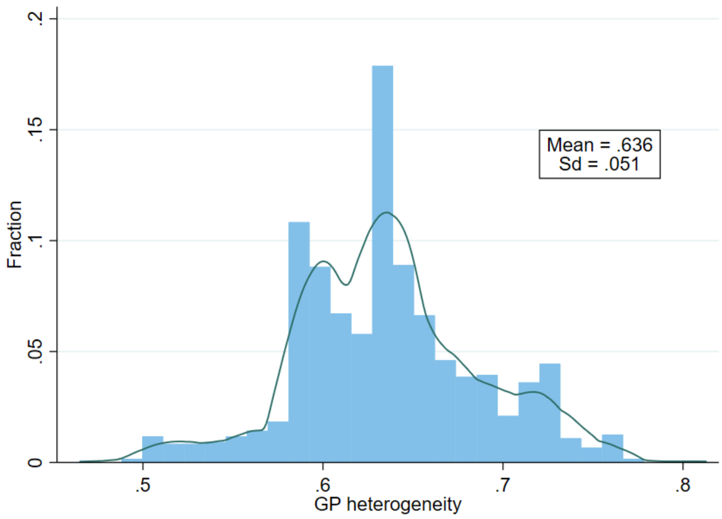


Table 1: Land quality and subsistence practices

	Gathering	Hunting	Fishing	Husbandry	Agriculture	Agriculture intensity
Land quality	-0.150	0.008	-0.256	-0.165	0.563	0.147
	(0.052)***	(0.047)	(0.043)***	(0.050)***	(0.075)***	(0.052)***
R^2	0.40	0.44	0.35	0.42	0.50	0.29
N	1,236	1,236	1,236	1,236	1,236	1,144
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Land quality is computed taking into account the most productive crop of each pixel. Standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

4 Econometric analysis

4.1 Baseline model

The baseline model is a linear equation with the Ethnographic Atlas societies as cross-sectional unit of observation:

$$y_s = \alpha + \beta GP_s^{het} + X_s'\psi + u_s \tag{1}$$

Where: y_s is the measure of state-centralisation in society s; GP_s^{het} is the measure of GP heterogeneity as averaged across all the pixels in an ethnic polygon; X_s is a vector of controls, which in the baseline specification includes mean rainfall and temperature, absolute latitude, and continent fixed effects; u_s are standard errors with arbitrary spatial correlation within a 200km radius from the polygon centroid.[21]

The model estimates a reduced form, inasmuch as we observe only potential, rather than actual, growing period heterogeneity. While the measure is arguably exogenous, claims of causality are limited by the cross-sectional nature of the data. A series of robustness tests gives, however, an interpretation of the results beyond the mere correlation.

Table 2 reports results from Equation 1 as estimated through Ordinary Least Square (OLS), ordered Probit and Logit. All the point estimates of β are negative and highly significant. The effect is extremely large. The conditional OLS estimate is almost equal to minus four, that is, roughly four standard deviations of the dependent variable. To put it differently, a passage from full homogeneity to full heterogeneity of the crop cycle translates into a full-scale political collapse: the passage from a large state to an acephalous society. Such type of abrupt change is obviously absent from the data, with GP^{het} ranging between 0.46 and 0.81. Marginal impacts are, however, substantial also in terms of standard deviations: a one standard deviation increase in crop cycle heterogeneity triggers a 0.2 standard deviation drop in jurisdictional hierarchy. Baseline estimates thus point to a strong, but reasonable, detrimental impact of growing period heterogeneity onto state centralisation.[22]

21. Figure A6 zooms the sample in Northwestern Africa, graphically illustrating the type of comparison implicit in Equation 1.

22. Another way to interpret the baseline marginal impact is to express it in its actual unit: days. For example, an increase of three months of common agricultural season ($-\Delta GP_s^{het} = 3/12 = 1/4 \sim 5\sigma_{GP_s^{het}}$) corresponds to an additional level of jurisdictional hierarchy.

Table 2: State centralisation and crop cycle heterogeneity - Baseline regressions

Dependent variable: Hierarchy levels above the local community				
	OLS	OLS	Logit	Probit
	(1)	(2)	(3)	(4)
GP heterogeneity	-3.167	-3.945	-8.920	-5.096
	(0.630)***	(0.683)***	(1.448)***	(0.819)***
	{0.779}***	{0.711}***	[1.816]***	[1.027]***
	{{0.812}}***	{{0.740}}***		
	{{{0.872}}}***	{{{0.742}}}***		
R^2	0.02	0.27	0.13	0.13
N	1,066	1,061	1,061	1,061
Mean dependent variable	.902	.906	.906	.906
Baseline controls	No	Yes	Yes	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Standard errors in square brackets are robust standard errors. Standard errors in single, double and triple curly brackets are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 50km, 100km and 200km, respectively. Standard errors in squared brackets are clustered at the regional level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

4.2 Robustness tests

A Ruling out alternative hypotheses

The baseline model substantiates a strong negative impact of the growing period heterogeneity on state-building. This section provides evidence that the observed correlation is not driven by some omitted variable related to both state centralisation and the variability of the agricultural calendar.

Table 3 reports results from Equation 1 when controlling sequentially and cumulatively for various geographic characteristics of the societies: ruggedness, river discharge, number of rivers, distance to coastlines, polygon area, and variability of temperature and precipitation patterns. The coefficients of these variables go in the expected direction (*e.g.* positive effect of river count on state centralisation), but the magnitude and significance of growing period heterogeneity barely change.

Table 4 controls for agricultural factors traditionally associated to state-building: the productive advantage of (storable) cereals over (perishable) roots and tubers, land quality, land circumscription, and the number of economically relevant crops. The coefficient of interest (β) remains always negative and significant.

A third set of possible confounders is given by trade variables. Crop cycle heterogeneity could, indeed, be associated also to mutual insurance, given the potential different exposure of crops to common shocks. Table 5 shows that when controlling for various proxies of trade,[23] results closely mimic baseline estimates.

Finally, Table 6 checks for a series of socio-economic variables that are closely connected to state centralisation. These measures come mostly from the Ethnographic Atlas and captures: cereal cultivation, dependence on agriculture, pastoralism, use of plough, and historic conflict. These results should be taken with a grain of salt, given the probable endogenous nature of most of these controls. Nevertheless, the point estimates of interest is always negative and statistically significant.

23. Trade covariates include: indexes of ecological fractionalisation and polarisation, the standard deviation of land quality, a measure of how much the most productive crop change across neighbouring areas, and an index of subsistence fractionalisation.

Table 3: State centralisation and crop cycle heterogeneity - Control for additional geographic factors

Dependent variable: Hierarchy levels above the local community

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GP heterogeneity	-3.931	-3.947	-3.959	-3.951	-4.075	-3.791	-3.951	-3.967
	(0.749)***	(0.747)***	(0.735)***	(0.743)***	(0.740)***	(0.744)***	(0.732)***	(0.734)***
Ruggedness	-0.025							0.013
	(0.036)							(0.038)
River discharge		-0.002						-0.052
		(0.024)						(0.017)***
River count			0.000					0.000
			(0.000)***					(0.000)***
Coast distance				-0.000				-0.000
				(0.000)				(0.000)
Area					0.133			0.143
					(0.023)***			(0.023)***
Rain stdev						0.005		0.008
						(0.003)**		(0.003)***
Temperature stdev							0.019	-0.008
							(0.154)	(0.165)
R^2	0.27	0.27	0.29	0.27	0.31	0.27	0.27	0.33
N	1,061	1,061	1,061	1,061	1,061	1,061	1,061	1,061
Mean dependent var.	.906	.906	.906	.906	.906	.906	.906	.906
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table 4: State centralisation and crop cycle heterogeneity - Control for agricultural factors
Dependent variable: Hierarchy levels above the local community

	(1)	(2)	(3)	(4)	(5)	(6)
GP heterogeneity	-2.077	-2.370	-3.806	-3.290	-1.813	-2.006
	(0.954)**	(0.950)**	(0.742)***	(0.795)***	(0.935)*	(0.924)**
Cereal advantage	0.143				1.946	1.589
	(0.041)***				(0.413)***	(0.382)***
Land quality		0.124			-1.955	-1.601
		(0.041)***			(0.442)***	(0.408)***
Land circumscription			0.148		0.136	0.209
			(0.050)***		(0.051)***	(0.066)***
Productive crops				0.015	0.030	0.030
				(0.007)**	(0.010)***	(0.010)***
R^2	0.28	0.28	0.28	0.27	0.31	0.38
N	1,061	1,061	1,061	1,061	1,061	1,061
Mean dependent variable	.906	.906	.906	.906	.906	.906
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	No	No	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, area, and the standard deviations of rainfall and temperature. Standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table 5: State centralisation and crop cycle heterogeneity - Control for trade proxies
Dependent variable: Hierarchy levels above the local community

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GP heterogeneity	-3.791	-3.915	-3.326	-3.954	-3.861	-3.021	-3.356
	(0.713)***	(0.729)***	(0.764)***	(0.780)***	(0.740)***	(0.792)***	(0.780)***
Ecological fractionalisation	0.529					1.002	0.614
	(0.138)***					(0.313)***	(0.315)*
Ecological polarisation		3.791				-6.606	-5.123
		(1.596)**				(3.600)*	(3.578)
Land quality variability			0.098			0.078	0.053
			(0.041)**			(0.041)*	(0.039)
Crop type heterogeneity				0.022		-0.295	-0.018
				(0.297)		(0.278)	(0.316)
Subsistence fractionalisation					-0.927	-0.931	-0.872
					(0.283)***	(0.275)***	(0.257)***
R^2	0.28	0.27	0.28	0.27	0.28	0.30	0.35
N	1,061	1,061	1,061	1,061	1,061	1,061	1,061
Mean dependent variable	.906	.906	.906	.906	.906	.906	.906
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	No	No	No	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, area, and the standard deviations of rainfall and temperature. Standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table 6: State centralisation and crop cycle heterogeneity - Control for socio-economic factors
Dependent variable: Hierarchy levels above the local community

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GP heterogeneity	-3.922	-2.963	-4.152	-2.789	-3.936	-2.353	-2.278
	(0.759)***	(0.703)***	(0.736)***	(0.617)***	(0.739)***	(0.582)***	(0.578)***
Cereal main	0.201					-0.150	-0.136
	(0.086)**					(0.083)*	(0.083)*
Agriculture dependence		0.107				0.109	0.121
		(0.021)***				(0.023)***	(0.020)***
Pastoralism			0.011			0.011	0.011
			(0.002)***			(0.002)***	(0.002)***
Plough				1.098		0.841	0.715
				(0.171)***		(0.175)***	(0.163)***
Conflict					0.000	0.000	0.000
					(0.001)	(0.001)	(0.001)
R^2	0.27	0.31	0.29	0.34	0.27	0.38	0.43
N	1,052	1,061	1,061	1,058	1,061	1,049	1,049
Mean dependent variable	.901	.906	.906	.908	.906	.904	.904
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	No	No	No	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, area, and the standard deviations of rainfall and temperature. Standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

B Additional sensitivity checks

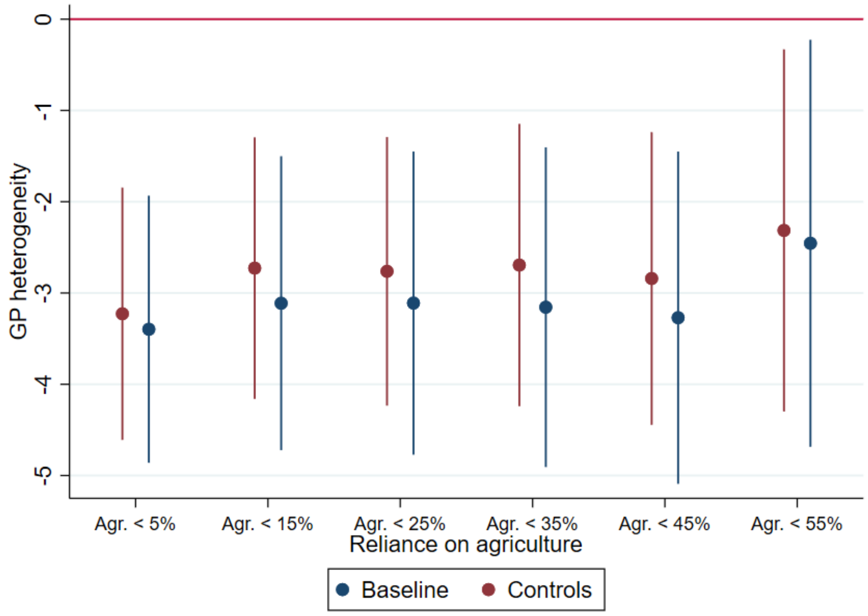
Two batteries of robustness tests are run. The first pertains sample-wide modifications whereby: (*i*) societies with low reliance on agriculture are excluded from the sample, inasmuch as the constraints of an heterogeneous agricultural calendar should be felt only by farming societies; (*ii*) each continent is dropped sequentially, given that estimates are based on within-continent variation; (*iii*) alternative geographical representations of the Ethnographic Atlas societies are employed.

Figure 7 shows that as we restrict the sample to societies with high reliance on agriculture, point estimates remain negative and significant. Similarly, when singularly excluding one continent at time, results are robust as shown in Figure 8. Table 7 reports, instead, estimates of Equation 1 when Ethnographic Atlas societies are represented as circles of varying radii (25km, 50km, and 100km) built around the centroid of the society.[24] Point estimates are highly significant and of a magnitude comparable to the baseline equation.

The second series of robustness tests concerns the definition of the main independent and dependent variables. Table 8 shows that growing period heterogeneity is also a good predictor of the extensive margins of state centralisation, that is, the presence of at least one level of jurisdictional hierarchy above the local community. Figure 9 checks robustness of the estimates to alternative pixel dimensions (0.083°, 0.25°, and 0.5°) and neighbourhood size (9, 25, 49). Results are negative and significant, albeit they decrease in absolute magnitude as the pixel and neighbourhood sizes are increased. As substantiated in Appendix B.1, this likely signals that what really hampered state extraction was a very localised form of crop cycle heterogeneity.

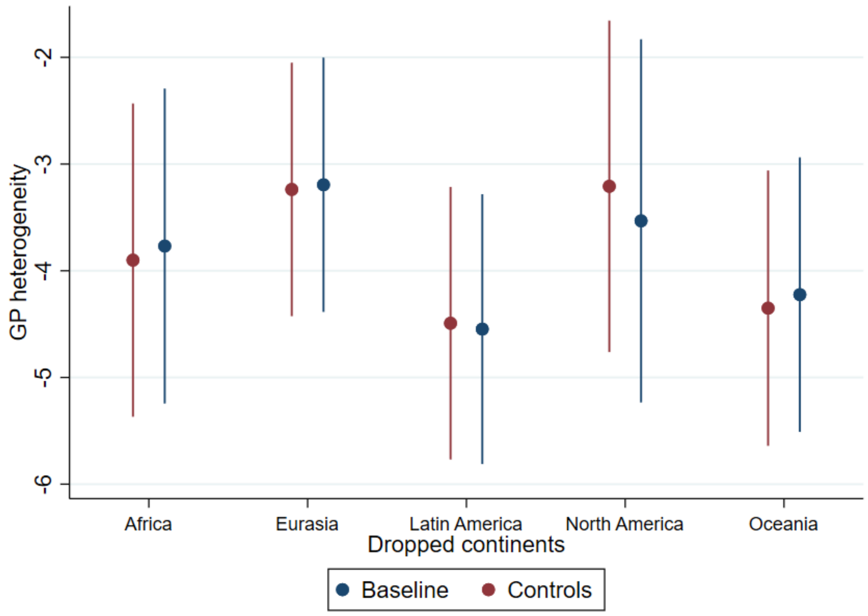
24. Figure A5 gives a visual representation of this alternative geographical representation.

Figure 7: Sample robustness - Drop societies with low dependence on agriculture



Notes: *Notes*: For the definition of controls see Table 7. The figure reports 90% confidence intervals.

Figure 8: Sample robustness - Drop continents



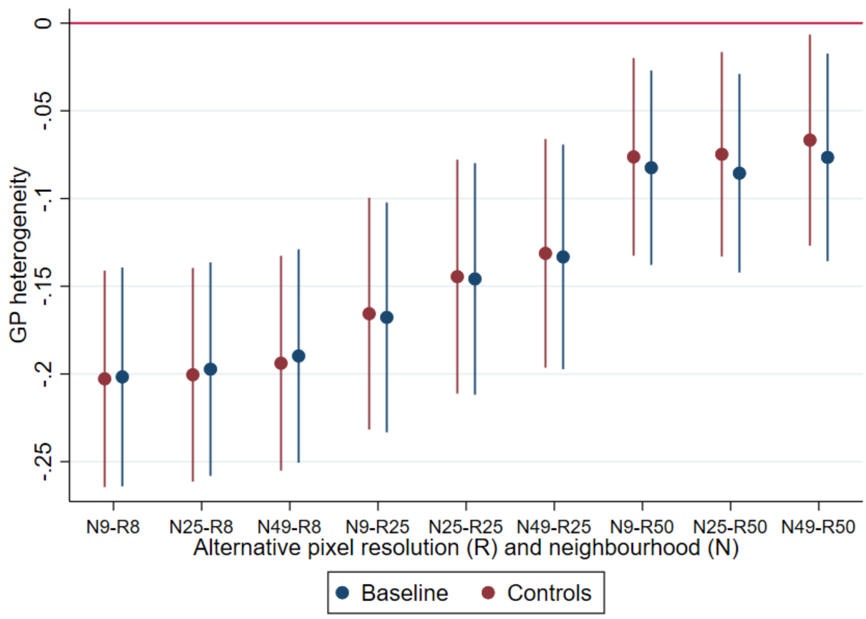
Notes: *Notes*: For the definition of controls see Table 7. The figure reports 90% confidence intervals.

Table 7: State centralisation and crop cycle heterogeneity - Alternative samples
Dependent variable: Hierarchy levels above the local community

	Buffer 25		Buffer 50		Buffer 100	
GP heterogeneity	-2.668	-2.701	-3.414	-3.466	-4.248	-4.332
	(0.715)***	(0.707)***	(0.781)***	(0.768)***	(0.842)***	(0.831)***
R^2	0.26	0.27	0.26	0.28	0.27	0.28
N	1,009	1,009	1,027	1,027	1,044	1,044
Mean dependent variable	.901	.901	.903	.903	.897	.897
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	Yes	No	Yes	No	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, area, and the standard deviations of rainfall and temperature. Standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Figure 9: State centralisation and crop cycle heterogeneity - Alternative pixel and neighbourhood sizes



Notes: For the definition of controls see Table 7. Treatment variables are standardised. The figure reports 90% confidence intervals.

Table 8: State centralisation and crop cycle heterogeneity - Alternative state centralisation measure

Dependent variable: Presence of hierarchy above the local community

	OLS		Logit		Probit	
GP heterogeneity	-1.420	-1.371	-7.476	-7.367	-4.490	-4.415
	(0.322)***	(0.329)***	(2.193)***	(2.115)***	(1.292)***	(1.243)***
R^2	0.24	0.26	0.18	0.21	0.18	0.20
N	1,061	1,061	1,061	1,061	1,061	1,061
Mean dependent variable	.906	.906	.906	.906	.906	.906
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	Yes	No	Yes	No	Yes

The table reports different probability model estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, area, and the standard deviations of rainfall and temperature. In OLS regressions standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. In Logit and Probit estimations standard errors are clustered at the regional level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

4.3 Mechanisms: Taxation

This section provides empirical evidence on one of the channels whereby crop cycle heterogeneity hampered state formation in the pre-industrial era. As already discussed at section 2, in places marked by a scattered agricultural calendar, taxation was more difficult to sustain. In highly centralised states, the collection of land taxes - which alongside corvée duties formed the bulk of state resources - was underpinned by impressive accounting systems, aptly designed for censusing and recording farming activity. Homogenous local environmental conditions - be them land fertility, climatic patterns, or crop cycles - made production more transparent and assessable, thus decreasing the chances for tax evasion (Mayshar et al. 2017). Conversely, when agricultural working schedules varied too much within the same region, oversight and categorisation of farming was often too complicated. For example, Qing rulers struggled in managing grain surcharges in provinces that had more heterogeneous grain collection costs, leaving their management to local authorities (Ch'ü 1962, ch. 8). Similarly, it has been noted how, in the case of Ancient Egypt, "the high transparency of farming helps explains why the Pharaohs were able to [...] siphon off a substantial share of the tax revenue" (Mayshar et al. 2017, p. 630). In short, the heterogeneity of crop cycles was ill-suited to the regularity needs of centralised tax collection, which, indeed, tended to follow a precise uniform schedule within each region.[25]

Subnational tax data is extremely difficult to retrieve for the late pre-industrial era. Partial evidence is available for 19^{th} century China and India. As for the first, there is information on agricultural tax revenues levied in each prefecture in 1820.[26] In the case of the Indian subcontinent, Dincecco et al. (2022) have assembled data on 1881 land tax revenues across districts ruled either directly or indirectly by the United Kingdom. The left panel of Figure 10 shows tax revenues per km^2 for Qing China and colonial India. On the right side of the figure, subnational districts are shaded by the average level of growing period heterogeneity: the inverse relationship between the latter and tax revenues stands out clearly.[27]

In order to test more formally the relationship between taxation and crop cycle

25. In Qing China (1644-1912), for example, taxes were collected in the second to fourth month, then paused between months 5 and 7 (farmers' busy season) and finished in months 8-11 (Ch'ü 1962, ch. 8).

26. The original data is collected in Liang, Fangzhong (2008). *Statistics on Chinese Historical Demography, Land, and Land Tax.* Beijing: Zhonghua Book Company. The version here used has been kindly shared with me by Ma (2022).

27. Note that for China, districts correspond to historical prefectures, whose borders at 1820 were retrieved from the China Historical Geographic Information System (ChGIS) website. For India, similar GIS data is unavailable, so that, while tax data is at the historical district level, growing period heterogeneity is at a slightly more aggregated level given by current administrative level-2 divisions.

Table 9: Land tax and crop cycle heterogeneity - Baseline regressions
Dependent variable: Land tax revenue per km^2 (19^{th} century)

	India	China	India	China	India	China
GP heterogeneity	-4.107	-5.824	-7.076	-7.397	-7.560	-5.866
	(0.740)***	(1.018)***	(1.072)***	(1.816)***	(1.223)***	(1.469)***
R^2	0.05	0.08	0.10	0.24	0.19	0.33
N	438	261	438	261	438	261
Baseline controls	No	No	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	No	Yes	Yes

The table report OLS estimates. The dependent variable is standardised to have mean zero and unit standard deviation. Baseline controls include: land quality, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, and the standard deviations of rainfall and temperature. Robust standard errors are reported in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

heterogeneity, I specify the following linear model:

$$tax_d^s = \alpha^s + \beta^s GP_d^{het} + X_d' \gamma^s + \varepsilon_d \quad (2)$$

Where: tax_d^s is land tax per km^2 in district d of society $s \in \{China, India\}$, standardised so as to have mean zero and unit standard deviation; GP_d^{het} is the index of crop cycle heterogeneity; X_d is a vector of controls, which in its baseline version includes land fertility, mean rainfall and temperature, and absolute latitude; ε_d are robust standard errors.[28]

The identification of the coefficients of interest (β^{China} and β^{India}) relies on the arguably exogenous nature of GP_d^{het}, which is based on potential, rather than realised, data. Moreover, with respect to Equation 1, here, the cross-sectional units of observation are inherently more comparable one another, belonging to the same political unit.

Table 9 displays OLS estimates from Equation 2. The first two columns report point estimates of β^s from univariate regressions, baseline controls are added in columns 3-4, and a further set of geographical covariates is cumulatively added at columns 5-6. The impact of growing period heterogeneity is always negative and highly significant. Marginal effects are also substantial, with one standard deviation increase in GP^{het} being associated to 0.29-0.37 (0.20-0.37) standard deviation drop in tax revenues across China (India).

In Appendix B.2, it is shown that the relationship is robust to: *(i)* the inclusion of additional agro-ecological, trade, and socio-economic covariates (*e.g.*

28. I follow Dincecco et al. (2022) in specifying standard errors as robust. As shown in Appendix B.2, estimates remain highly significant even when employing more conservative error specifications, whereby ε_d is clustered regionally or is adjusted for spatial correlation.

population density and distance from capital); *(ii)* the use of regional fixed effects (administrative level-1 units for the Indian sample, historical provinces for the Chinese sample); *(iii)* the sequential exclusion of these regions; *(iv)* alternative definitions of the dependent variable (expressed in logs) and of growing period heterogeneity (computed with varying pixel and neighbourhood sizes); *(v)* the control of overall state capacity, as proxied by the number of Mogul sites for India and the non-independent nature of prefectures in China. The latter sensitivity test is particular relevant inasmuch as it pins down more precisely the taxation channel, ruling out that crop cycle heterogeneity simply proxies for underlying levels of state presence.

Finally, this section concludes with some tentative evidence on one of the precise ways whereby scattered agricultural calendars lowered taxation. Beyond resource legibility, crop cycle heterogeneity might have hampered taxation directly in terms of resource appropriation. When crops in adjoining fields are *harvested* at different moments in time, tax collectors shall be mobilised at various periods of the year. Travelling to a same area several times per year is obviously more costly and complicated than to collect the bulk of yearly production all at once. In short, places with heterogeneous harvest dates should have been more difficult to tax.

Mimicking the baseline measure of crop cycle heterogeneity and employing the same underlying GAEZ data, I thus define a variable capturing harvest heterogeneity. Define H_p as the day when the most productive crop of pixel p is harvested. Hence, harvest heterogeneity is defined as:

$$H_p^{het} = \frac{1}{|N_p|} \sum_{k \in N_p} |H_p - H_k|$$

Table 10 reports OLS estimates from a version of Equation 2 substituting GP_d^{het} with the average value of H_p^{het} across pixels belonging to Indian and Chinese districts. The first two columns only include baseline controls, the successive two pairs add cumulatively further geographical controls and regional fixed effects. All coefficients are negative and mostly significant. The only exception being the point estimate relative to the Chinese sample employing the full set of control and fixed effects, which becomes marginally insignificant.

Overall the exercise further gives credit to the idea that crop cycle heterogeneity had a causal impact on state-building, lowering state capacity by, *inter alia*, making taxation more arduous.

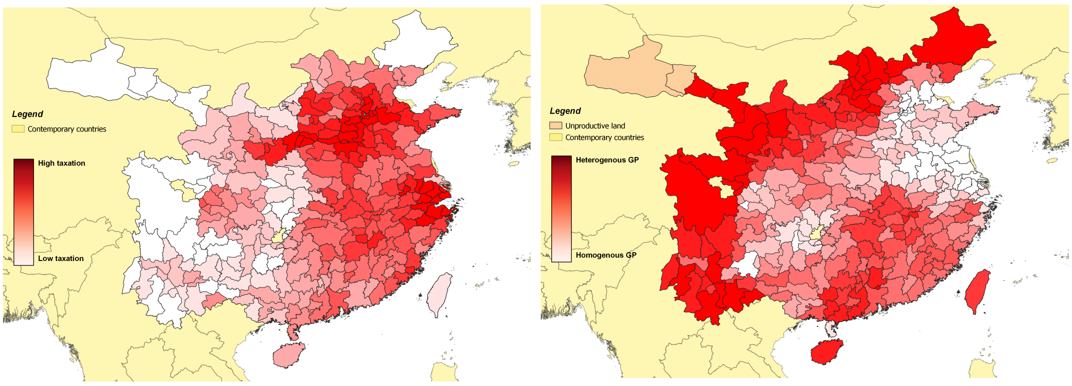
Table 10: Land tax and crop cycle heterogeneity - Harvest date heterogeneity
Dependent variable: Land tax revenue per km^2 (19^{th} century)

	India	China	India	China	India	China
Harvest heterogeneity	-5.042	-1.605	-4.349	-1.633	-3.293	-1.353
	(0.901)***	(0.804)**	(0.948)***	(0.725)**	(1.449)**	(0.837)
R^2	0.07	0.20	0.15	0.30	0.25	0.46
N	438	261	438	261	438	261
Mean dependent variable	.42	.013	.42	.013	.42	.013
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	Yes	Yes	Yes	Yes
Region fixed effects	No	No	No	No	Yes	Yes

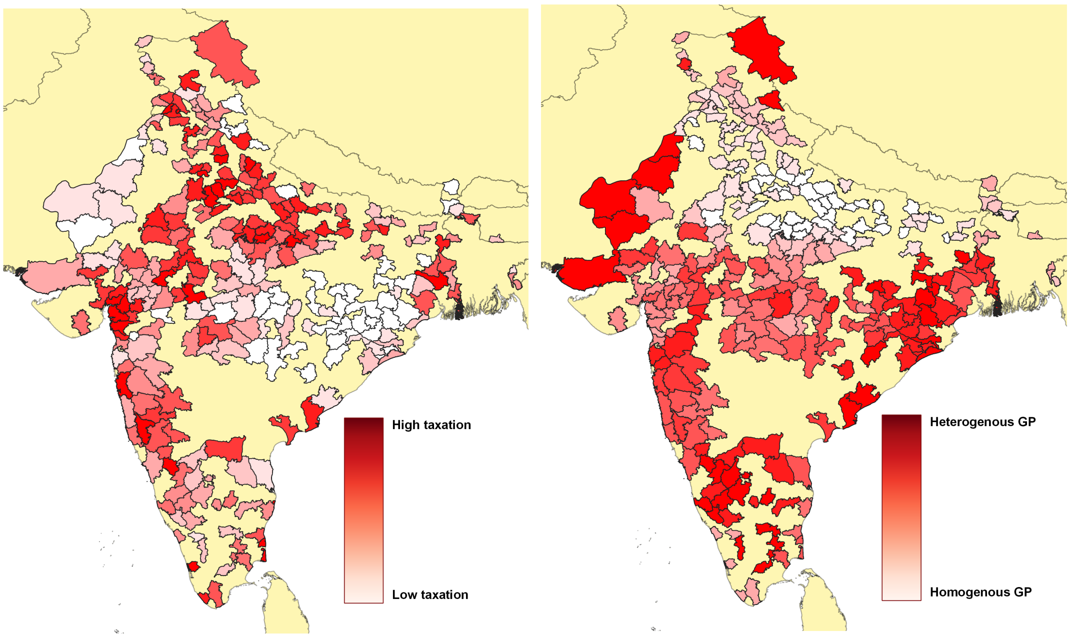
The table report OLS estimates. The dependent variable is standardised to have mean zero and unit standard deviation. Baseline controls include: land quality, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, and the standard deviations of rainfall and temperature. Robust standard errors are reported in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Figure 10: Land tax and growing period heterogeneity in Qing China (1820) and British India (1881)

(a) China



(b) India



5 Conclusions

Contrary to popular perceptions, centralised governments have been rather fragile constructions for much of the pre-industrial era. Their emergence and consolidation was highly dependent on a set of very particular agro-ecological conditions, which allowed for a stable and regular tax base. As shown in this paper, the homogeneity of the agricultural calendar was among such constraints. Where crops in adjoining fields followed different growth cycles, the accounting and eventual taxation of agricultural output required a more prolonged and extended effort, which was often beyond the capabilities of most agrarian polities. Moreover, in areas without a common agricultural season, coordination of corvée activities and mass recruitment in the army were effectively prevented, further limiting the venues to state formation. Finally, heterogeneous agricultural calendars translated into different social arrangements regulating the tempo and rhythm of communal life. The growing cycle of the main staple crop mandated, indeed, for particular working schedules and religious practices. Their fragmentation, as mandated by heterogeneous farming calendars, might have thus represented a further barrier to the emergence of centralised authorities, which have historically relied on a rather uniform social base. In short, centralised governments could not emerge in agro-ecological settings defying their homogenisation attempts.

Overall, this paper sheds lights on one of the most daunting question of comparative history and social sciences at large: the origin of the state. It does so by putting forward and testing empirically a relatively neglected dimension of state-building: the constraint represented by heterogeneous crop cycles. Its appreciation enrich our understanding on the uneven historical development of state institutions, providing further insights into the different development trajectories of the various areas of the world.

References

Adams, Robert McC. 2007. "The limits of state power on the Mesopotamian Plain." *Cuneiform Digital Library Bulletin* 1:1–3.

Algaze, Guillermo. 2001. "Initial social complexity in southwestern Asia: the Mesopotamian advantage." *Current Anthropology* 42 (2): 199–233.

———. 2009. *Ancient Mesopotamia at the dawn of civilization.* University of Chicago Press.

Allen, Robert C., Mattia C. Bertazzini, and Leander Heldring. 2020. "The Economic Origins of Government." Working paper, Kellogg School of Management, Northwestern University.

Baum, Richard. 2004. "Ritual and Rationality: Religious Roots of the Bureaucratic State in Ancient China." In *The Early State: Its Alternatives and Analogues,* edited by Leonid E. Grinin, Robert L. Carneiro, Dimitri M. Bondarenko, Nikolay N. Kradin, and Andrey V. Korotayev. Volgograd: 'Uchitel' Publishing House.

Bellwood, Peter. 2004. *First farmers: the origins of agricultural societies.* John Wiley & Sons.

Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and markets: The advantage of an early start." *Journal of Economic growth* 7 (4): 347–369.

Borcan, Oana, Ola Olsson, and Louis Putterman. 2018. "State history and economic development: evidence from six millennia." *Journal of Economic Growth* 23 (1): 1–40.

Bowles, Samuel, Eric Alden Smith, and Monique Borgerhoff Mulder. 2010. "The emergence and persistence of inequality in premodern societies: introduction to the special section." *Current Anthropology* 51 (1): 7–17.

Butzer, Karl W. 2012. "Collapse, environment, and society." *Proceedings of the National Academy of Sciences* 109 (10): 3632–3639.

Carballo, David M, Paul Roscoe, and Gary M. Feinman. 2014. "Cooperation and collective action in the cultural evolution of complex societies." *Journal of Archaeological Method and Theory* 21 (1): 98–133.

Carmona, Salvador, and Mahmoud Ezzamel. 2007. "Accounting and accountability in ancient civilizations: Mesopotamia and ancient Egypt." *Accounting, Auditing & Accountability Journal.*

Carneiro, Robert L. 1970. "A Theory of the Origin of the State: Traditional theories of state origins are considered and rejected in favor of a new ecological hypothesis." *science* 169 (3947): 733–738.

Ch'ü, T'ung-tsu. 1962. *Local Government in China under the Ch'ing.*

Chatfield, Charlotte. 1953. "Food Composition Tables for International Use." *FAO Nutritional Studies,* no. 3.

Claessen, Henri J. M. 2004. "Was the State Inevitable?" In *The Early State: Its Alternatives and Analogues,* edited by Leonid E. Grinin, Robert L. Carneiro, Dimitri M. Bondarenko, Nikolay N. Kradin, and Andrey V. Korotayev. Volgograd: 'Uchitel' Publishing House.

Clastres, Pierre. 1987. *Society against the state: Essays in political anthropology.* New York: Zone Books.

Diamond, Jared. 1997. *Guns, Germs and Steel: The fates of human societies.* New York: WW Norton.

Dickson, D. Bruce. 1987. "Circumscription by anthropogenic environmental destruction: an expansion of Carneiro's (1970) theory of the origin of the state." *American Antiquity* 52 (4): 709–716.

Dincecco, Mark, James Fenske, Anil Menon, and Shivaji Mukherjee. 2022. "Pre-colonial warfare and long-run development in india." *The Economic Journal* 132 (643): 981–1010.

Dusinberre, Elspeth R. M. 2013. *Empire, authority, and autonomy in Achaemenid Anatolia.* Cambridge University Press.

Enke, Benjamin. 2019. "Kinship, cooperation, and the evolution of moral systems." *The Quarterly Journal of Economics* 134 (2): 953–1019.

FAO. 2001. *Food balance sheets, a handbook.* Food / Agriculture Organisation of the United Nations.

Fenske, James. 2014. "Ecology, trade, and states in pre-colonial Africa." *Journal of the European Economic Association* 12 (3): 612–640.

Flannery, Kent V. 1998. "Order, Legitimacy, and Wealth in Ancient Egypt and Mesopotamia." In *Archaic States,* edited by G. M. Feinman and J. Marcus, 15–58. Santa Fe: School of American Research Press.

Galor, Oded, and Ömer Özak. 2016. "The agricultural origins of time preference." *American Economic Review* 106 (10): 3064–3103.

Gat, Azar. 2006. *War in human civilization.* Oxford University Press.

Gennaioli, Nicola, and Ilia Rainer. 2007. "The modern impact of precolonial centralization in Africa." *Journal of Economic Growth* 12 (3): 185–234.

Graeber, David, and David Wengrow. 2021. *The dawn of everything: A new history of humanity.* Penguin UK.

Grinin, Leonid E. 2004. "The Early State and Its Analogues: A Comparative Analysis." In *The Early State: Its Alternatives and Analogues,* edited by Leonid E. Grinin, Robert L. Carneiro, Dimitri M. Bondarenko, Nikolay N. Kradin, and Andrey V. Korotayev. Volgograd: 'Uchitel' Publishing House.

Gurven, Michael, Monique Borgerhoff Mulder, Paul L. Hooper, Hillard Kaplan, Robert Quinlan, Rebecca Sear, Eric Schniter, et al. 2010. "Domestication alone does not lead to inequality: intergenerational wealth transmission among horticulturalists." *Current Anthropology* 51 (1): 49–64.

Halstead, Paul. 1989. "The economy has a normal surplus: economic stability and social change among early farming communities of Thessaly, Greece." In *Bad year economics: cultural responses to risk and uncertainty,* edited by P. Halstead and J. O'Shea, 68–80. Cambridge University Press Cambridge.

Harari, Yuval Noah. 2014. *Sapiens: A brief history of humankind.* Random House.

Henrich, Joseph, and Robert Boyd. 2008. "Division of labor, economic specialization, and the evolution of social stratification." *Current Anthropology* 49 (4): 715–724.

Ikenberry, G. John. 2011. "The future of the liberal world order: Internationalism after America." *Foreign affairs* 90 (3): 56–68.

Kelly, Robert L. 2007. *The foraging spectrum: Diversity in hunter-gatherer lifeways.* Smithsonian Inst Press.

Litina, Anastasia. 2014. "The geographical origins of early state formation." Working paper, CREA Discussion Paper.

Lozny, Ludomir R. 2004. "Transition to Statehood in Central Europe." Chap. 12 in *The Early State: Its Alternatives and Analogues,* edited by Leonid E. Grinin, Robert L. Carneiro, Dimitri M. Bondarenko, Nikolay N. Kradin, and Andrey V. Korotayev, 278–287. Volgograd: 'Uchitel' Publishing House.

Ma, Chicheng. 2022. "Classicism and Modern Growth: The Shadow of the Sages." *Available at SSRN 4288048.*

Mayoral, Laura, and Ola Olsson. 2019. "Pharaoh's Cage: Environmental Circumscription and Appropriability in Early State Development." Working paper.

Mayshar, Joram, Omer Moav, and Zvika Neeman. 2017. "Geography, transparency, and institutions." *American Political Science Review* 111 (3): 622–636.

Mayshar, Joram, Omer Moav, and Luigi Pascali. 2022. "The Origin of the State: Land Productivity or Appropriability?" *Journal of Political Economy* 130 (4): 1091–1144.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. "Pre-colonial ethnic institutions and contemporary African development." *Econometrica* 81 (1): 113–152.

Mulder, Monique Borgerhoff, Ila Fazzio, William Irons, Richard L. McElreath, Samuel Bowles, Adrian Bell, Tom Hertz, and Leela Hazzah. 2010. "Pastoralism and wealth inequality: revisiting an old question." *Current Anthropology* 51 (1): 35–48.

Mullerson, Rein. 1993. "The Continuity and Succession of States, by Reference to the former USSR and Yugoslavia." *International & Comparative Law Quarterly* 42 (3): 473–493.

Murdock, George P. 1967. *Ethnographic atlas.* University of Pittsburgh Press.

Nunn, Nathan, and Nancy Qian. 2011. "The potato's contribution to population and urbanization: evidence from a historical experiment." *The quarterly journal of economics* 126 (2): 593–650.

Possehl, Gregory L. 1998. "Sociocultural complexity without the state: The Indus civilisation." In *Archaic States,* edited by G. M. Feinman and J. Marcus, 199–260. Santa Fe: School of American Research Press.

Rahmstorf, Lorenz. 2012. "Control mechanisms in Mesopotamia, the Indus Valley, the Aegean and Central Europe, c. 2600–2000 BC, and the question of social power in early complex societies." *Beyond elites: alternatives to hierarchical systems in modelling social formations* 215:311–326.

Redmond, Elsa M., and Charles S. Spencer. 2012. "Chiefdoms at the threshold: The competitive origins of the primary state." *Journal of Anthropological Archaeology* 31 (1): 22–37.

Sánchez De La Sierra, Raúl. 2020. "On the origins of the state: Stationary bandits and taxation in eastern congo." *Journal of Political Economy* 128 (1): 32–74.

Sato, Tsuneo. 1990. "Tokugawa Japan." Chap. Tokugawa Villages and Agriculture, edited by Chie Nakane and Shinzaburo Oishi, 37–80. University of Tokyo Press.

Schönholzer, David. 2020. "The Origin of the Incentive Compatible State: Environmental Circumscription." Working paper.

Scott, James C. 2009. *The art of not being governed.* Yale University Press.

———. 2017. *Against the grain.* Yale University Press.

Shaw, Malcolm N. 2014. *International Law.* 7th. Cambridge University Press.

Shenk, Mary K., Monique Borgerhoff Mulder, Jan Beise, Gregory Clark, William Irons, Donna Leonetti, Bobbi S Low, et al. 2010. "Intergenerational wealth transmission among agriculturalists: foundations of agrarian inequality." *Current Anthropology* 51 (1): 65–83.

Smith, Eric Alden, Kim Hill, Frank W. Marlowe, David Nolin, Polly Wiessner, Michael Gurven, Samuel Bowles, Monique Borgerhoff Mulder, Tom Hertz, and Adrian Bell. 2010. "Wealth transmission and inequality among hunter-gatherers." *Current Anthropology* 51 (1): 19–34.

Smith, Eric Alden, Monique Borgerhoff Mulder, Samuel Bowles, Michael Gurven, Tom Hertz, and Mary K. Shenk. 2010. "Production systems, inheritance, and inequality in premodern societies: conclusions." *Current anthropology* 51 (1): 85–94.

Spencer, Charles S. 2010. "Territorial expansion and primary state formation." *Proceedings of the National Academy of Sciences* 107 (16): 7119–7126.

Tedeschi, Gianluca. 2021. *Two sides of the same coin: Co-evolution of kin ties and institutions.* Technical report. Mimeo, University of Nottingham.

Trigger, Bruce G. 2003. *Understanding early civilizations: a comparative study.* Cambridge University Press.

Webster, David. 1975. "Warfare and the Evolution of the State: A Reconsideration." *American Antiquity* 40 (4): 464–470.

Online Appendix to

Crop cycles and hierarchy: the agro-ecological origins of the state

Matteo Sestito[†]

†. Aix-Marseille University, CNRS, EHESS, Centrale Marseille, AMSE, France.

Contents

A Theories of state-building

Traditionally, theories on (pristine) state-formation have been divided in two types: voluntaristic and coercive. The first emphasise bottom-up process of state formation, whereby the latter developed because it performed some common-interest function. Coercive theories of state-building, instead, stress factors related to the ability of central authorities to tax and control subservient masses. I abandon this categorisation to eschew difficult, perhaps unanswerable, philosophical questions on the voluntaristic or coercive nature of the processes historically linked to state-building. Therefore, below I list several macro social phenomena that have been proposed as causes of the process of (early) state-building. None of them shall be singularly understood as necessary, let alone sufficient, and their effect clearly depends on other structural factors such as: subsistence systems, ideological constructs, and population size.[1]

Conflict is one of the most cited forces behind archaic state-building (Redmond and Spencer 2012, Webster 1975). In the first place, communal defence and security can be thought as public goods, making their central administration more efficient. Throughout history looming external threats have, indeed, repeatedly prompted loose tribal confederations to unite under a sole banner (*e.g.* Gallic tribes against Julius Caesar, Israeli tribes against Ammonites), paving the way to the development of centralised systems of control (Gat 2006). Moreover, taking the perspective of the attacker, territorial expansion by itself implies, beyond a certain range, the delegation of power and hence the establishment of bureaucratic structures of command (Spencer 2010). This mechanism is even more compelling when vanquished populations have no possibility to escape and are thus forced into a relation of subservience (Carneiro 1970, Dickson 1987). Econometric evidence on this latter channel comes from Schönholzer (2020), who find that pristine state formation is associated to land circumscription, measured as the differential in land productivity between a zone and the neighbouring areas.

The latter analysis, modelling land circumscription rather than warfare per se, is also coherent with a second factor traditionally associated to state-building, namely, the easiness whereby some central authority can extol taxes. Hence, land circumscription, by decreasing the possibilities of outmigration, makes population control and tax collection easier. Beyond the work of Schönholzer (2020), empirical evidence is available for ancient Egypt, where state power is correlated to positive

1. For a thorough discussion of how social stratification unfolds in different subsistence systems, the reader is referred to: Bowles et al. (2010), Smith, Hill, et al. (2010), Smith, Mulder, et al. (2010), Mulder et al. (2010), Gurven et al. (2010), Shenk et al. (2010). For the role of ideology and religion in shaping the early development of the Chinese state, see Baum (2004). Finally, for a theoretical model on the interactions between population size, technological innovation and labour & social stratification, the reader is referred to Henrich and Boyd (2008).

productive shocks in the core Egyptian territories and negative agricultural shocks in its periphery (Mayoral and Olsson 2019).
Alongside migration possibilities, another element positively influencing the tax base of (would-be) states is the presence of patchy, regular, and appropriable resources (Smith, Mulder, et al. 2010). The presence of storable and predictable agricultural surpluses is particularly relevant in discussions on the consolidation of the first states and is often considered almost a necessary condition for their emergence (Scott 2017). Empirical evidence on this point comes from Mayshar et al. (2022), who find that archaic and pre-modern states were more centralised where the production of storable crops such as cereals enjoyed advantages over the cultivation of more perishable roots and tubers.
Crop cycles represent a new, relatively neglected, agro-ecological constraint on the fiscal capacity of pre-industrial polities. Importantly, as shown in section 4, the impact of heterogeneous growing seasons on political centralisation is largely orthogonal to land circumscription and the so-called cereal advantage.

Other theories tend to emphasise not much the coercive side of state-building, but rather its benefits in terms of provision and maintenance of public infrastructures such as temples and irrigation networks. The most prominent among these theories is Karl Wittfogel's 'hydraulic hypothesis', which explained archaic state formation as a result of the collective effort geared towards irrigation. Archaeological discoveries have cast some doubts on the latter, illustrating how often states preceded large-scale centralised irrigation structures (Carneiro 1970, Carballo et al. 2014). Yet, recent econometric analysis by Allen et al. (2020) indicates that, at least in Mesopotamia, state-building in its early days responded to the collective action problems related to the construction and maintenance of irrigation canals. Heterogeneous agricultural growing seasons impaired the ability of central authorities to raise substantial labour forces in the provision of public goods, ultimately blocking a further venue to state formation.

A fourth set of causes connected to the consolidation of states, concerns those mechanisms based on economic exchange. Trade, for example, figures prominently in the rise of the city-state system of 4^{th} millennium BCE Mesopotamia (Algaze 2001, 2009). By fostering economic growth and labour specialisation, trade is generally associated to deepening social inequality, thus possibly resulting into more politically stratified societies. Evidence of these mechanisms is not limited to full-blown state polities, but include, for example, foragers of the north-western Pacific Coast (Kelly 2007) as well as the rural communities of archaic Thessaly (Halstead 1989). More generally, economic exchange between unrelated communities can strengthen the position of the elite by either giving them a public function (*e.g.* protection of trade, construction of roads) or directly enriching them (*e.g.* taxation on trade). The state trajectory of the Yoruba Oyo polity in present-day Nigeria and Benin, as well as that of many other pre-colonial African kingdoms, seems, indeed, to follow this scheme, whereby central authorities relied extensively on the control of trade activities. Fenske (2014) provides related econometric evidence on pre-industrial African state-building; Litina (2014) and Tedeschi (2021) uncover, instead, broad correlations between trade and state emergence at the global level.

B Mechanisms

B.1 Localised vs. aggregate growing period heterogeneity

When discussing the venues whereby crop cycle heterogeneity hampered state formation, the stress has been put on the channels whereby a heterogeneous agricultural calendar constrained centralised resource extraction. The basic underlying idea is that, in virtue of their very heterogeneity, working schedules differing within a short range were difficult to be monitored, organised and eventually exploited. However, it might well be argued that a scattered agricultural calendar could facilitate resource collection. It might have allowed states to maintain lean bureaucracies, with tax collectors and state administrators activated sequentially in each region as the main regional agricultural season came to end. These rotating administrations - sometimes consisting of the whole court as in late Shang China or Carolingian France -[2] did, indeed, provide an efficient and relatively inexpensive way to administer territories. Yet, their success was ultimately dependent on local conditions, including homogenous crop cycle *within* the region where they were deployed. While crop cycle diversity across regions might have even encouraged state-building, when felt it in a more localised form across fields belonging to the same region, it constrained the ability of local magistrates to census production and extol resources (be them in the form of labour services or financial transfers).

The discussion thus boils down to understanding whether the spatial unit of analysis here employed is well suited to capture a localised form of crop cycle heterogeneity. It shall be noted that GAEZ data is quite fine grained, coming at a resolution of 0.083°, that is, in cells measuring 9 × 9 kilometres at the equator and slightly less as we move towards the poles. A first way to gauge the validity of this cell size is to compare it to the dimension of administrative units in state polities. If GAEZ cells approximate these units, then crop cycle heterogeneity across neighbouring cells may actually incentivise state-building by allowing a rotating schedule for tax collection and labour mobilisation. While it is difficult to establish the typical size of an administrative unit throughout the pre-industrial

2. For Shang China (1600-1046 BCE) see Trigger (2003, p. 109); for Carolingian France (751-1120 CE) see: https://www.britannica.com/biography/Charlemagne/Court-and-administration.

era, it can be quite confidently said that they were not smaller than GAEZ cells. On average, the 1293 administrative units of local government in Jiaqing's China (1760-1820) measured about 3235 km^2 (Ch'ü 1962).[3] Casting this area into a cell, we would have a square measuring 57 km by side, that is, roughly 6-7 GAEZ pixels. Even in more crowded environments hosting city-state systems, distances between (rival) administrative centres ranged between 7 and 67 kilometres.[4] Therefore, the baseline measure of growing period heterogeneity is unlikely to capture advantages inherent into rotating administrative schemes.

Results at Figure 9 are consistent with these observations, whereby lower pixel sizes are associated to stronger marginal impacts on state centralisation.[5] Moreover, when confronting measures of growing period heterogeneity at the baseline cell size level (0.083°) with more aggregate definitions (computed with interpolated GAEZ data), only the former has a negative significant impact. Figure A1 reports OLS estimates of a version of Equation 1 including an additional term, $\widetilde{GP}_s^{het}$: the measure of growing period heterogeneity built employing pixels of greater sizes (0.25°, 0.5°, and 1°). All regressions include the vectors of baseline and geographical controls. Regardless of the size of the cell neighbourhood across which heterogeneity is computed, only the localised version of GP_s^{het} has a significantly detrimental impact on state centralisation. When computed at greater cell size (ranging from 784 to 12.321 km^2), aggregated crop cycle heterogeneity, conditional on its localised version, has a positive, if mostly insignificant, impact. This perhaps signals the advantages inherent in having a variegated production cycle across macro-regions.

3. The area was computed on the 1820 prefecture polygons assembled by ChGIS.

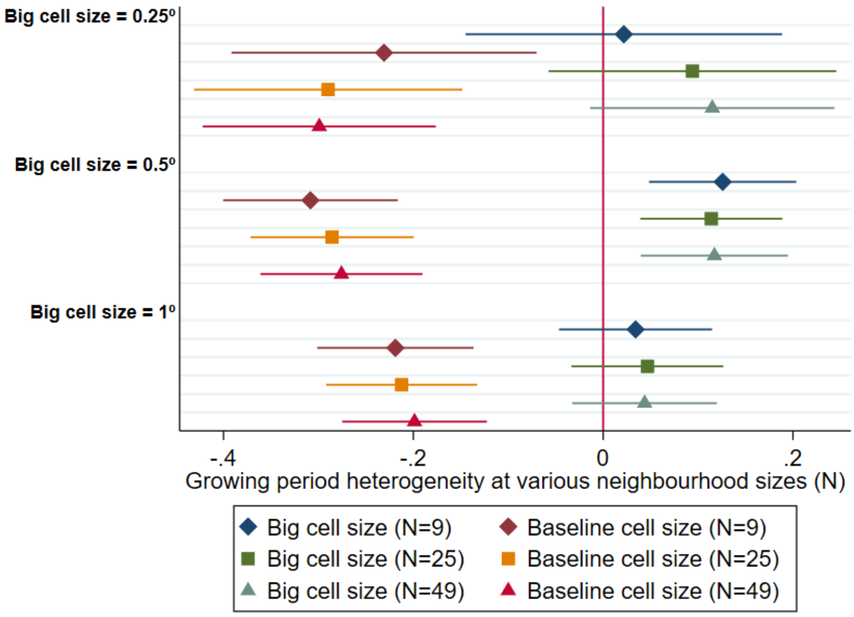
4. Sumer Mesopotamia is an outlier on this regard, with rival cities being often visible from each others' walls (Trigger 2003, p. 100). Yet, each city complex measured probably about 40 km in diameter.

5. Recall that:

$$GP_s^{het} = \frac{1}{|N_s|} \sum_{p \in N_s} \frac{1}{|N_p|} \sum_{k \in N_p} 1 - \frac{|GP_p \cap GP_k|}{365}$$

with p indicating a pixel of a given dimension and $|N_p|$ the size of the pixel neighbourhood.

Figure A1: Localised vs. aggregated growing period heterogeneity



textitNotes: The figure reports OLS regressions with state centralisation as dependent variable and *GP* heterogeneity at two different aggregation level as independent variables of interest. Dependent and treatment variable have been standardised. All regressions include baseline and geographic controls, for whose definition see Table 9. The figure reports 90% confidence intervals.

B.2 Taxation

Table A1: Land tax and crop cycle heterogeneity - Additional controls
Dependent variable: Land tax revenue per km^2 (19^{th} century)

	India	China	India	China	India	China
GP heterogeneity	-9.734	-3.686	-6.252	-6.273	-5.334	-2.805
	(1.674)***	(1.149)***	(1.322)***	(1.466)***	(1.348)***	(1.349)**
R^2	0.21	0.40	0.20	0.33	0.24	0.39
N	438	261	438	261	392	261
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Agro-ecological controls	Yes	Yes	No	No	No	No
Trade controls	No	No	Yes	Yes	No	No
Socio-economic controls	No	No	No	No	Yes	Yes

The table report OLS estimates. The dependent variable is standardised to have mean zero and unit standard deviation. Baseline controls include: land quality, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, and the standard deviations of rainfall and temperature. Agro-ecological controls include: the productive advantage of cereals over roots & tubers, land circumscription, and number of productive crops. Trade controls include: an index of ecological fractionalisation and polarisation, the standard deviation of land quality, and spatial heterogeneity of the most productive crop. Socio-economic controls include: log population density in 1800 for China and 1850 for India, distance from capital (Kolkata & Beijing), and pre-colonial conflict exposure for the Indian sample. Robust standard errors are reported in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

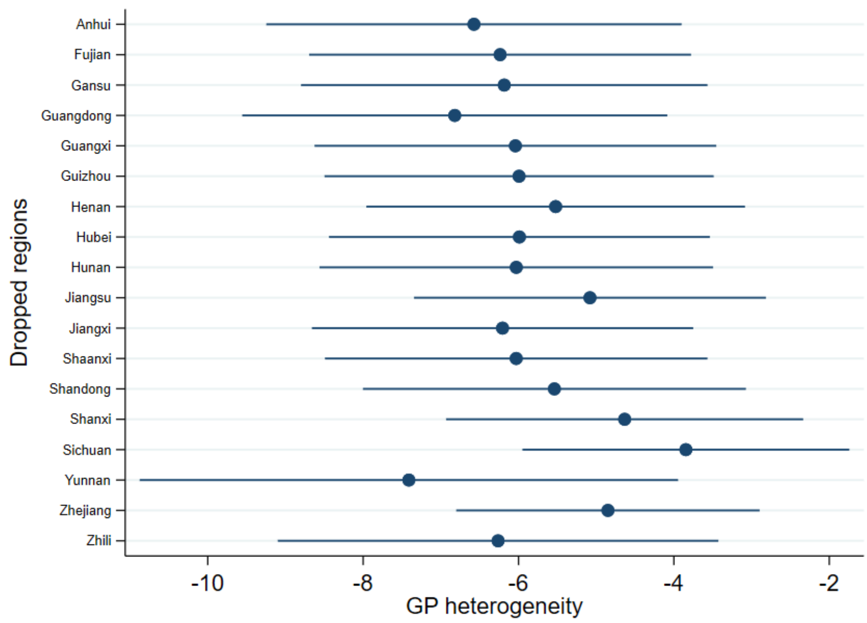
Table A2: Land tax and crop cycle heterogeneity - Robustness
Dependent variable: Land tax revenue per km^2 (19^{th} century)

	Revenues in levels		Log revenues			
	India	China	India	China	India	China
GP heterogeneity	-8.480	-5.768	-11.474	-2.430	-11.405	-6.256
	(2.181)***	(1.783)***	(1.944)***	(2.696)	(2.418)***	(3.083)**
R^2	0.26	0.48	0.37	0.62	0.55	0.72
N	438	261	438	261	438	261
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	No	No	Yes	Yes

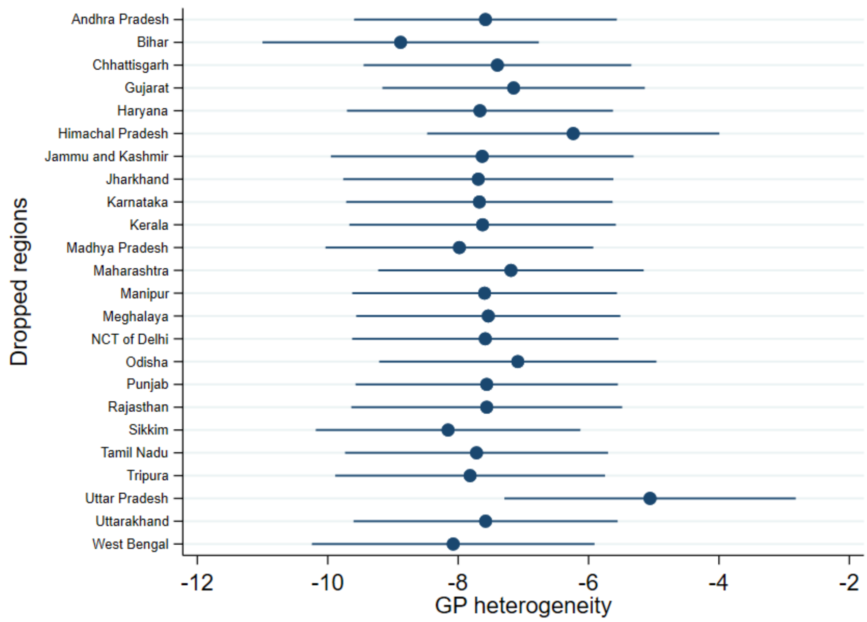
The table report OLS estimates. The dependent variable in levels is standardised to have mean zero and unit standard deviation. Baseline controls include: land quality, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, and the standard deviations of rainfall and temperature. Region fixed effects correspond to dummies for administrative level-1 regions in India and historical provinces in China. Robust standard errors are reported in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Figure A2: Land tax and growing period heterogeneity - Drop regions

(a) China



(b) India



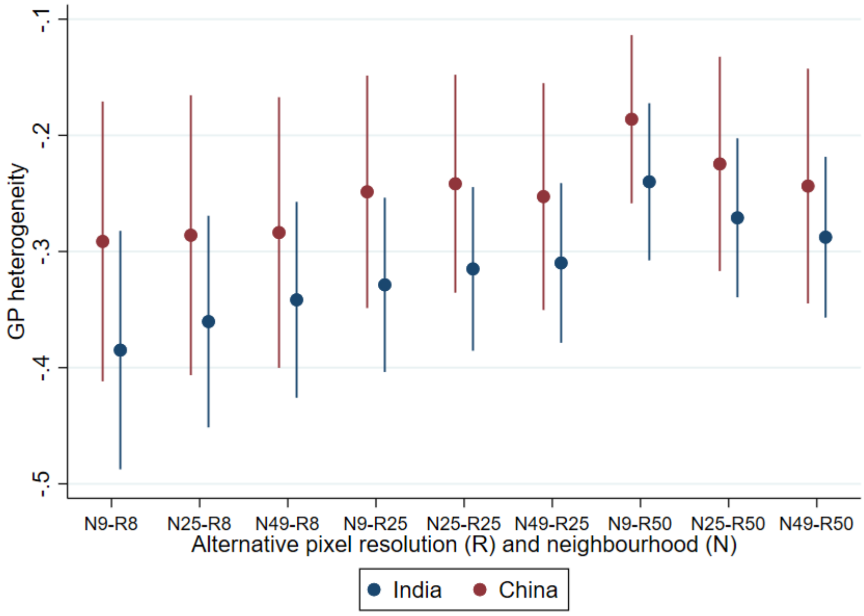
Notes: The figure reports 90% confidence intervals.

Table A3: Land tax and crop cycle heterogeneity - Control for state presence
Dependent variable: Land tax revenue per km^2 (19^{th} century)

	India	China	India	China	India	China
GP heterogeneity	-6.902	-6.982	-7.477	-5.576		
	(1.078)***	(1.837)***	(1.218)***	(1.509)***		
State presence	0.165	0.241	0.144	0.210	0.214	0.209
	(0.100)*	(0.084)***	(0.100)	(0.090)**	(0.103)**	(0.093)**
Harvest heterogeneity					-4.933	-1.228
					(0.958)***	(0.737)*
R^2	0.11	0.26	0.19	0.33	0.16	0.30
N	438	260	438	260	438	260
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	Yes	Yes	Yes	Yes

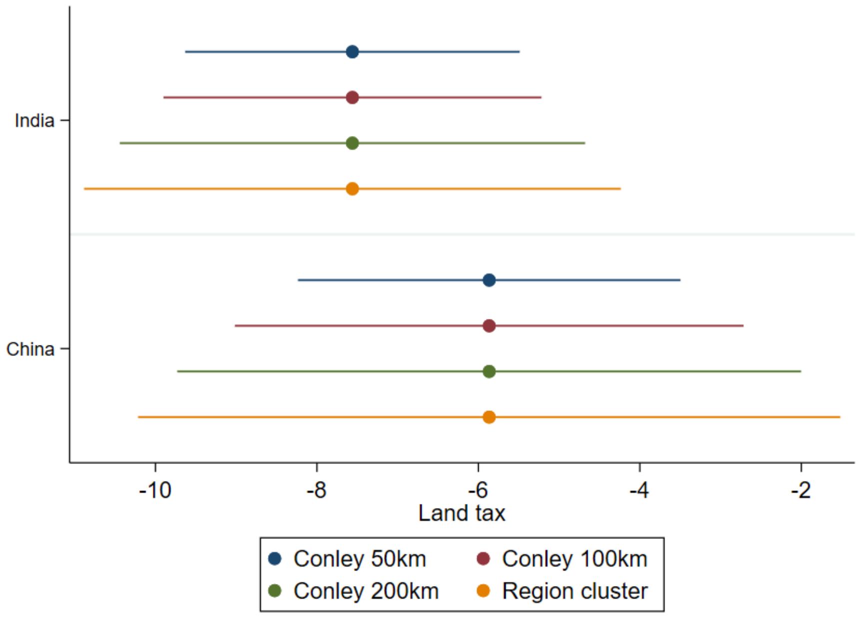
The table report OLS estimates. The dependent variable is standardised to have mean zero and unit standard deviation. Baseline controls include: land quality, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, and the standard deviations of rainfall and temperature. State presence is proxied by the number of Mogul sites for India & by the non-independent nature of the prefecture in China. Robust standard errors are reported in parenthesis. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Figure A3: Land tax and growing period heterogeneity - Alternative pixel and neighbourhood sizes



Notes: Dependent and treatment variable have been standardised. All regressions include baseline and geographic controls, for whose definition see Table 9. The figure reports 90% confidence intervals.

Figure A4: Land tax and growing period heterogeneity - Alternative standard error specifications



Notes: All regressions include baseline and geographic controls, for whose definition see Table 9. The figure reports 90% confidence intervals.

C Additional Maps

Figure A5: Ethnographic Atlas societies - Buffer 100 km representation

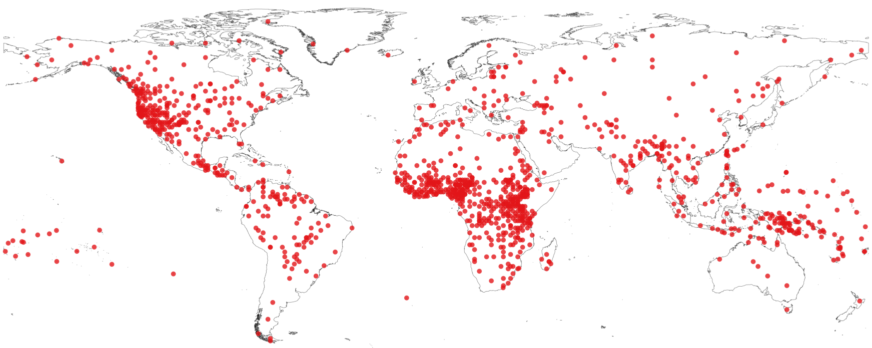
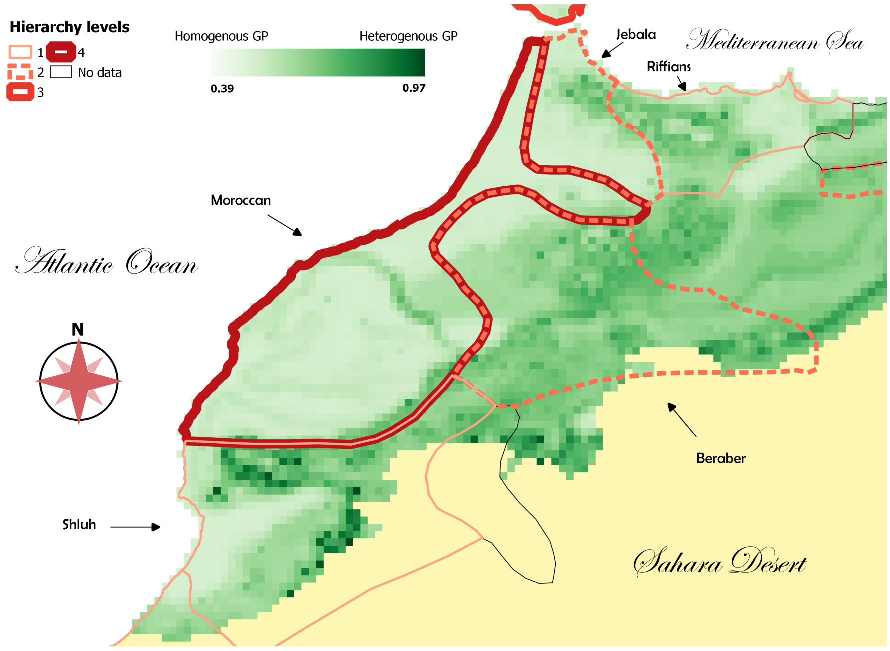


Figure A6: Ethnographic Atlas societies - Zoom on northwestern Africa



D Additional Tables

Table A1: Crop calories

Crop	Calories per 100g	Crop	Calories per 100g
Buckwheat	330	Phaseolus bean	341
Cabbage	19	Pigeonpea	343
Carrot	38	Rapeseed	494
Chickpea	358	Silage maize	356
Cotton	253	Soybean	335
Cowpea	342	Spring barley	332
Dry pea	346	Spring rye	319
Dryland rice	357	Spring wheat	334
Flax	534	Sugarbeet	70
Foxtail millet	343	Sunflower	308
Gram	345	Sweet potato	92
Greater yam	101	Temperate maize	356
Groundnut	567	Temperate sorghum	343
Highland maize	356	Tomato	17
Highland sorghum	343	Wetland rice	357
Lowland maize	356	White potato	67
Lowland sorghum	343	White yam	101
Oat	385	Winter barley	332
Onion	31	Winter rye	319
Pearl millet	348	Winter wheat	334

Sources: Chatfield (1953), FAO (2001), Galor and Özak (2016).

E Further material

Table A5: Folklore and crop cycle heterogeneity

Dependent variable: Log concept count in folklore motifs

	Annual	Annual	Off season	Off season
GP heterogeneity	-0.591	-0.539	0.478	0.463
	(0.201)***	(0.198)***	(0.276)*	(0.268)*
R^2	0.49	0.49	0.23	0.25
N	1,165	1,165	1,165	1,165
Baseline controls	Yes	Yes	Yes	Yes
Geographic controls	No	Yes	No	Yes

The table report OLS estimates. Baseline controls include: continent fixed effects, absolute latitude, and mean rainfall and temperature. Geographic controls include: ruggedness, river flow discharge, number of rivers, distance to coast, log area, and the standard deviations of rainfall and temperature. All regressions control also for the total number of motifs and year of their first publication. Standard errors are adjusted for spatial correlation using Conley's (1999) method with a distance cut-off of 200km. *, **, and *** indicate significance at the 10%, 5%, and 1% levels.

Table A6: Pairwise raw correlation coefficients

	GP het.	Lat.	Rain avg	Temp. avg	Rugged.	River count	Coast dist.	Cereal adv.	Land quality	Land circ.	Prod. crops	Ecol. fract.	Land qual. var.	Crop type het.
GP heterogeneity	1.00													
Latitude (abs)	0.41	1.00												
Rain avg	-0.42	-0.47	1.00											
Temperature avg	-0.43	-0.89	0.40	1.00										
Ruggedness	0.14	0.31	0.02	-0.36	1.00									
River count	-0.00	0.04	0.00	-0.04	-0.05	1.00								
Coast distance	0.13	-0.08	-0.33	-0.01	-0.22	0.08	1.00							
Cereal advantage	-0.46	0.08	0.02	-0.02	0.03	0.01	0.02	1.00						
Land quality	-0.50	-0.03	0.08	0.08	-0.01	0.01	0.03	0.99	1.00					
Land circumscription	-0.18	-0.07	0.34	0.13	0.02	-0.03	-0.25	0.12	0.12	1.00				
Productive crops	-0.38	-0.25	0.16	0.28	-0.03	-0.01	0.09	0.70	0.76	0.05	1.00			
Ecological fractionalisation	0.16	0.32	-0.30	-0.39	0.26	0.08	-0.02	0.11	0.08	-0.19	0.08	1.00		
Land quality variability	-0.03	0.40	-0.16	-0.35	0.40	0.09	-0.04	0.44	0.38	0.06	0.22	0.33	1.00	
Crop type heterogeneity	0.24	0.22	-0.06	-0.35	0.57	-0.04	-0.13	-0.18	-0.21	-0.08	-0.18	0.24	0.26	1.00